А. Н. ТИХОНОВ
А. С. ЛЕОНОВ
А. Г. ЯГОЛА

НЕЛИНЕЙНЫЕ НЕКОРРЕКТНЫЕ ЗАДАЧИ

А. Н. ТИХОНОВ
А. С. ЛЕОНОВ
А. Г. ЯГОЛА

НЕЛИНЕЙНЫЕ НЕКОРРЕКТНЫЕ ЗАДАЧИ

МОСКВА
НАУКА • ФИЗМАТЛИТ
1995

ББК 22.193
Т46
УДК 519.6

Федеральная целевая программа книгоиздания России

Т и х о н о в А. Н., Л е о н о в А. С., Я г о л а А. Г. Нелинейные некорректные задачи.—М.: Наука. Физматлит, 1995.—312 с.—ISBN 5-02-014582-3.

Описаны эффективные численные алгоритмы для решения некорректно поставленных задач. Приведены регуляризующие алгоритмы, предназначенные для решения плохо обусловленных систем линейных уравнений, интегральных уравнений Фредгольма, в том числе несовместных, одномерных и двумерных уравнений типа свертки.

Для научных работников, специализирующихся в области решения обратных задач, задач математического проектирования, обработки наблюдений и т. п.

Ил. 43. Библиогр. 242 назв.

Р е ц е н з е н т доктор физико-математических наук *А. В. Крянев*

$\text{Т}\frac{1602120000-021}{053(02)-95}$ Без объявл.

ISBN 5-02-014582-3

© А. Н. Тихонов, А. С. Леонов, А. Г. Ягола, 1995

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Книга посвящена вариационным методам решения нелинейных, некорректно поставленных задач. Некорректно поставленные задачи часто встречаются в самых разнообразных приложениях (астрофизике, геофизике, диагностике плазмы, компьютерной томографии, оптимальном проектировании технических систем, оптимальном планировании, оптимальном управлении различными объектами и др.). С математической точки зрения эти задачи формулируются в форме операторных уравнений первого рода, задач минимизации функционалов, задач вычисления значений неограниченных операторов, вариационных неравенств и пр. В отличие от корректно поставленных (корректных по Адамару [229]) задач, решение которых существует, единственно и устойчиво зависит от данных задачи в заданном функциональном пространстве, некорректно поставленные задачи не могут быть решены с помощью классических методов вычислительной математики. Соответствующие вычислительные процедуры оказываются для них расходящимися.

Для преодоления этой трудности был предложен ряд новых методов, позволяющих получать устойчивые приближенные решения некорректно поставленных задач [84—88, 109—112, 189—202]. К ним относится и метод регуляризации с использованием сглаживающего функционала. На их основе можно построить серию вычислительных алгоритмов, различающихся такими характеристиками, как скорость, порядок точности получаемого приближения и т. д. Методы решения некорректных задач (в том числе метод регуляризации с использованием сглаживающего функционала) получили значительное развитие в 60—80-е годы. Далеко не полный список работ по этой тематике приводится в данной книге. Особо детальное исследование метод регуляризации получил для линейных некорректных задач. Это отражено, в частности, в монографиях [10, 19, 20, 27, 28, 53, 54, 71, 72, 88, 110, 112, 142, 160, 164, 187, 188, 202, 206, 209, 214, 230, 234].

Вместе с тем алгоритмы решения нелинейных некорректных задач еще не достаточно исследованы или, точнее говоря, изучены несколько односторонне. В основном развитие получили алгоритмы, основанные на априорном выборе параметра регуляризации. Алгоритмы же, связанные с апостериорным выбором этого параметра, для нелинейных задач общего вида недостаточно освещены в монографиях по некорректно поставленным задачам [19, 20, 32, 142, 164, 202, 206]. В немалой степени по этой причине появилась данная книга.

В книге излагаются теория и практика применения вариационных регуляризирующих алгоритмов решения нелинейных некорректных задач с апостериорным выбором параметра регуляризации. При этом не затрагивается ряд других важных подходов в теории нелинейных некорректных задач таких, например, как итеративная регуляризация [19, 20, 32, 203], решение вариационных неравенств [6, 20], решение операторных уравнений с монотонными операторами [6], статистический подход к некорректным задачам [209, 214].

Для удобства чтения в книге имеется вводная глава (при ссылках нумеруемая как нулевая), в которой дан основной справочный материал по некоторым разделам топологии, функционального анализа, теории функций и линейной алгебры.

Глава 1 посвящена теории вариационных методов решения некорректных экстремальных задач. В § 1, 2 ставятся основные экстремальные задачи (минимизации по функционалу и по аргументу), приводятся примеры их корректной и некорректной постановки и излагаются основные предположения. В § 3 дается устойчивый по отношению к возмущениям данных метод решения задач минимизации по функционалу, который существенно используется в дальнейшем. В § 4—6 изложен вспомогательный материал. В них развивается аппарат исследования алгоритмов устойчивой минимизации по аргументу. Сами же алгоритмы — обобщенные принципы невязки, квазирешений и сглаживающего функционала для экстремальных задач — формулируются и изучаются соответственно в § 7—9. В § 10 исследуются исключительные случаи, когда алгоритмы формально дают нулевое значение параметра регуляризации. В этом случае алгоритмы модифицируются к более простому виду. Подробному обсуждению конкретных условий применимости алгоритмов посвящен § 11, где излагаются удобные для практической проверки требования на данные задачи. В § 12 исследуется связь алгоритмов обобщенного принципа невязки и обобщенного принципа квазирешений с другими вариационными алгоритмами решения некорректных экстремальных задач — методом невязки и методом квазирешений. Важный вопрос о точности изучаемых алгоритмов (при определенной априорной информации о решении) рассматривается в § 13. Там показано, что алгоритмы обобщенных принципов невязки, квазирешений и сглаживающего функционала имеют оптимальный порядок точности. В § 14 разбираются вопросы численной реализации алгоритмов.

В гл. 2 алгоритмы обобщенных принципов невязки, квазирешений и сглаживающего функционала применяются для решения нелинейных операторных уравнений вариационным методом. Дается постановка задачи, связь ее с вариационными задачами из гл. 1 и методика ее решения (§ 1). Затем рассматриваются алгоритмы, специально адаптированные для случая операторных уравнений (§ 2), в частности для случая разрешимых операторных уравнений (§ 3). В § 4—6 предпринимается исследование этих алгоритмов, по целям близкое к проведенному в § 10—12 гл. 1, но отличающееся по содержанию. Вопросу о сравнении алгоритмов с априорным и апостериорным выбором параметра регуляризации посвящен § 7. В нем отмечаются преимущества и недостатки каждого из этих типов алгоритмов. Глава 2 завершается кратким рассмотрением задач с избыточной информацией, т. е. таких задач, в которых имеется достаточно большой набор приближенных реализаций данных. Эти реализации могут отличаться друг от друга по точности. Обрабатывая определенным образом эти реализации, можно выделить наиболее точные и использовать их в дальнейшем в алгоритмах обобщенных принципов невязки, квазирешений и сглаживающего функционала.

В гл. 3 излагается теория конечномерных модификаций обобщенных принципов невязки, квазирешений и сглаживающего функционала. Дается схема конечномерной аппроксимации нелинейных некорректных задач (§ 1) и затем для решения получаемой конечномерной задачи строятся указанные конечномерные алгоритмы (§ 2—4). В § 5 приводятся конкретные примеры конечномерной аппроксимации нелинейных задач, для решения которых можно применять алгоритмы с апостериорным выбором параметра регуляризации.

В ряде случаев функция, представляющая искомое решение некорректной задачи, может быть непрерывной всюду, кроме не более чем счетного числа точек разрыва первого рода с неизвестным положением. При этом возникает проблема кусочно равномерной регуляризации. Она заключается в построении таких приближенных решений, которые сходились бы к точному равномерно на любом отрезке, не содержащем точек разрыва точного решения. Один из возможных подходов к этой проблеме (другие подходы см. в [9, 67, 73, 81, 206]) дается в гл. 4. Он основан на использовании функций с ограниченной вариацией. В § 1 приводятся вспомогательные сведения об этих функциях. В § 2 ставится задача о кусочно равномерной регуляризации и даются алгоритмы ее решения на основе материала из гл. 2, 3. Затем в § 3 приводятся примеры конкретных задач, для кусочно равномерной регуляризации которых применимы эти алгоритмы. В заключительном § 4 дается обоснование методике численной реализации алгоритмов кусочно равномерной регуляризации.

Предложенные и исследованные в гл. 2, 3 алгоритмы можно применять и для решения линейных задач. В частности, это относится и к решению вырожденных и плохо обусловленных систем линейных алгебраических уравнений. Общие алгоритмы решения некорректных задач применялись для систем линейных уравнений во многих работах [25, 28, 45, 55, 103, 105, 118, 154, 163, 167, 194, 197, 200, 201, 208, 227, 232, 241].

Основной целью было получение устойчивого приближения к решению (нормальному решению, нормальному псевдорешению) системы. При этом, однако, возникает также вопрос и о точности приближения. Проиллюстрируем это на примере метода регуляризации. Использование метода регуляризации для решения систем линейных алгебраических уравнений имеет одну особенность. Приближенное решение, получаемое с помощью этого метода по приближенным данным задачи, имеет оптимальный порядок точности, если система с точными данными разрешима. Если же последняя несовместна и ищется ее нормальное псевдорешение, то метод регуляризации, вообще говоря, может и не дать оптимального порядка точности приближения. Следует при этом отметить, что судить о разрешимости точной системы по ее приближенным данным в общем случае не представляется возможным.

Оказывается, что теория обобщенного принципа невязки для решения нелинейных некорректных задач, развитая в гл. 1, 2, может помочь преодолеть эти трудности. На ее основе можно сконструировать метод решения систем линейных уравнений с приближенными данными, точность которого не зависит от разрешимости или неразрешимости точной системы. Этот метод — метод минимальной псевдообратной матрицы (МПМ) — развивается в гл. 5.

В § 1 гл. 5 дается постановка задачи и далее (§ 2—4) формулируется и подробно исследуется метод МПМ. Этот метод обладает рядом важных оптимальных свойств, которые изучаются в § 5. Методика численной реализации метода МПМ рассматривается в § 6. Важной особенностью является тот факт, что по числу затрачиваемых на реализацию операций метод МПМ практически не отличается от других известных алгоритмов устойчивого решения систем линейных уравнений.

В гл. 6 приводятся некоторые примеры практических задач, для эффективного решения которых авторами были использованы алгоритмы с апостериорным выбором параметра регуляризации. Таковы, например, задачи колебательной спектроскопии, оптимизации режима лекарственной терапии, оптимального математического проектирования высокотемпературных излучателей, восстановления локальных профилей линий поглощения по наблюдаемым в спектре Ар-звезд и др.

В книге принята система ссылок на различные утверждения, имеющая форму $a.\ b.\ c.$ Здесь a — номер главы, где помещается данное утверждение; b — номер параграфа; c — порядковый номер утверждения данного типа. При ссылке внутри гл. a допускается неполный номер вида $b.\ c$, при ссылке внутри гл. a, § b — номер вида c.

Книга может быть полезной для студентов, аспирантов и научных работников, интересующихся нелинейными некорректными задачами.

ВВОДНАЯ ГЛАВА

НЕКОТОРЫЕ ВСПОМОГАТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ ТОПОЛОГИИ, ФУНКЦИОНАЛЬНОГО АНАЛИЗА И ЛИНЕЙНОЙ АЛГЕБРЫ

Современные методы решения некорректно поставленных задач завоевывают все большую популярность. Их все чаще используют при решении прикладных задач не только специалисты-математики, но и ученые, работающие в других областях науки, инженеры и др. В связи с этим для удобства чтения в книгу введена данная глава. Она содержит определения некоторых основных понятий из топологии, функционального анализа, теории функций, линейной алгебры, а также ряд касающихся их математических утверждений, которые непосредственно используются в книге. Даются также литературные источники, откуда можно почерпнуть более детальную информацию. При этом предполагается, что читатель имеет некоторую начальную подготовку в области математического анализа и линейной алгебры.

§ 1. Топологические пространства

Результаты, которые приводятся в данной книге, удобно излагать в общей форме в терминах теории топологических пространств. Приведем соответствующие основные понятия и факты этой теории.

Определение 1. Говорят, что на некотором множестве X задана *топология*, если в X определена система τ его подмножеств, которая удовлетворяет трем условиям: а) пустое множество $\varnothing$ и само X принадлежат семейству τ; б) объединение любого числа множеств из τ является снова множеством из τ; в) пересечение конечного числа множеств из τ также представляет собой множество из τ.

Семейство множеств τ называют *топологией* в X, а само множество X с заданной на нем топологией τ называют *топологическим пространством* и иногда обозначают символом (X, τ). Все множества T, входящие в τ, называются *открытыми*, а их дополнения $X\setminus T$ — *замкнутыми*.

Определение 2. *Окрестностью точки x топологического пространства* (X, τ) называется любое множество $T\subset\tau$, содержащее x. *Окрестностью множества* G $(G\subset X)$ называется всякое множество $T\subset\tau$, содержащее G.

Определение 3. Семейство $\mathscr{T}(G)$ окрестностей множества G $(G \subset X)$ называется *базой окрестностей* этого множества, если каждая окрестность множества G содержит некоторую окрестность из семейства $\mathscr{T}(G)$. В частности, множество G может состоять из одной точки.

Определение 4. Множество $\bar{G}$ называется *замыканием множества* G $(G \subset X)$, если оно представляет собой пересечение всех замкнутых множеств, содержащих G.

Определение 5. *Множество* G *плотно в* X, если $\bar{G}=X$. Топологическое пространство X, содержащее счетное плотное в X множество, называется *сепарабельным*.

Определение 6. *Топология* t *сильнее топологии* τ, *если* $\tau \subset t$. Ясно, что на X $(X \neq \varnothing)$ всегда существует *слабейшая топология*, состоящая из двух элементов: $\varnothing$ и X, а также *сильнейшая топология*, состоящая из всех подмножеств множества X (см. [93, с. 83]).

Определение 7. Топологическое пространство X называется *хаусдорфовым*, если всякие две различные его точки имеют непересекающиеся окрестности.

Определение 8. Пусть $\{x_n\}_{n=1}^{\infty} \subset X$ — последовательность элементов топологического пространства (X, τ). Говорят, что $\{x_n\}$ *сходится к точке* $x \in X$ *в топологии* τ (τ-*сходится*), если для любой окрестности T точки x найдется номер $N=N(T)$ такой, что $x_n \in T$ для всех $n \geqslant N$. При этом элемент x называется *пределом последовательности* $\{x_n\}$ *в топологии* τ.

Определение 9. Последовательность $\{x_n\}$ τ-*сходится к множеству* G $(G \subset X)$, если для любой окрестности T множества G найдется такой номер $N=N(T)$, что $x_n \in T$ для всех $n \geqslant N$.

Определение 10. Точка x называется *предельной точкой последовательности* $\{x_n\} \subset X$ *в топологии* τ (τ-*предельной точкой*), если для любой окрестности T точки x и любого номера m найдется такой номер n_m $(n_m > m)$, что $x_{n_m} \in T$.

Вообще говоря, в топологическом пространстве предельная точка последовательности может не являться пределом никакой последовательности (см., например, [32, с 61]).

Теорема 1. *В хаусдорфовом пространстве никакая последовательность не может иметь двух различных пределов.*

В принципе в топологическом пространстве (X, τ) можно ввести более общий вид сходимости, чем сходимость последовательностей. Эта так называемая *сходимость направлений* (или *сходимость по Муру — Смиту*) подробно описана в [90, гл. 1, § 2]. В этой книге она не используется.

Определение 11. Последовательность $\{x_n\}$, принадлежащая множеству D из топологического пространства (X, τ), называется τ-*секвенциально компактной в* D, если существует хотя бы одна ее подпоследовательность $\{x_{n_k}\}$, которая τ-сходится к некоторой точке $x \in D$.

Определение 12. Множество $D \subset X$ называется τ-*секвенциально компактным*, если любая последовательность $\{x_n\} \subset D$ является τ-секвенциально компактной.

В топологических пространствах имеются и другие виды компактности, кроме τ-секвенциальной компактности (τ-счетная компактность, τ-компактность). Подробно соответствующие понятия изложены в [90]. Для целей данной книги будет достаточно использовать понятие τ-секвенциальной компактности множества. Отметим также, что в случае метрических и банаховых пространств (см. ниже) все три понятия компактности оказываются эквивалентными.

Пусть D — множество из топологического пространства (X, τ).

Определение 13. Если каждому элементу $x \in D$ ставится в соответствие единственное число $J(x) \in \mathbf{R}$, то говорят, что *на D задан функционал $J(x)$*.

Определение 14. Функционал J, определенный на множестве D, называется *ограниченным снизу*, если существует такая константа M, что для любого x $(x \in D)$ выполнено неравенство $J(x) \geqslant M$.

Условие ограниченности снизу на множестве D функционала J эквивалентно конечности числа $J^* \equiv \inf\{J(x)\colon\ x \in D\}$.

Определение 15. Функционал $J(x)$, определенный на D, называется *τ-секвенциально полунепрерывным снизу (сверху) в точке $x \in D$*, если для любой последовательности $\{x_n\} \subset D$, τ-сходящейся к точке x, справедливо неравенство $\varliminf_{n \to \infty} J(x_n) \geqslant J(x)$ $(\varlimsup_{n \to \infty} J(x_n) \leqslant J(x))$.

Функционал $J(x)$ называется *τ-секвенциально полунепрерывным снизу (сверху) на множестве D*, если он τ-секвенциально полунепрерывен снизу (сверху) в каждой точке этого множества.

Функционал $J(x)$ называется *τ-секвенциально непрерывным в точке $x \in D$ (на множестве D)*, если он одновременно обладает свойствами τ-секвенциальной полунепрерывности снизу и сверху в точке x (на множестве D).

Определение 16. Последовательность элементов $\{x_n\}$ $(\{x_n\} \subset D)$ называется *минимизирующей для функционала $J(x)$ на множестве D*, если $\lim_{n \to \infty} J(x_n) = J^* = \inf\{J(x)\colon\ x \in D\}$.

Предположим, что функционал $J(x)$ определен на множестве D топологического пространства (X, τ). Тогда справедлив аналог известной теоремы Вейерштрасса.

Теорема 2 [32]. *Пусть множество D τ-секвенциально компактно, а функционал $J(x)$ τ-секвенциально полунепрерывен снизу на D. Тогда конечна точная нижняя грань $J^* = \inf\{J(x)\colon x \in D\}$ и множество $X_* = \{x \in D\colon\ J(x) = J^*\}$, содержащее элементы, которые реализуют эту точную нижнюю грань, не пусто. Множество X_* также τ-секвенциально компактно. Любая минимизирующая последовательность $\{x_n\}$ для функционала $J(x)$ на D τ-секвенциально компактна и τ-сходится к X_*.*

Условия этой теоремы достаточно удобны для проверки во многих конкретных случаях (см., например, § 3 и [32]).

Пусть теперь заданы топологические пространства (X, τ), (Y, t) и множества $D \subset X$, $R \subset Y$.

Определение 17. Если каждому элементу $x \in D$ соответствует единственный элемент $y = Ax \in R$, то говорят, что *задан оператор A, действующий из X в Y, с областью определения D и областью значений R.*

Символически это записывают в форме $A: X \to Y$. Для области определения и области значений оператора A приняты обозначения $D(A)=D$, $R(A)=R$.

Определение 18. Оператор $A: X \to Y$ называется *секвенциально непрерывным в точке* $x \in D(A)$, если для любой последовательности $\{x_n\} \subset D(A)$, τ-сходящейся к x, соответствующая последовательность $\{Ax_n\} \subset R(A)$ t-сходится к точке $Ax \in R(A)$.

Оператор *A секвенциально непрерывен на множестве G* ($G \subset D(A)$), если он секвенциально непрерывен в каждой его точке.

Подробное изложение теории топологических пространств можно найти, например, в книгах [90, 93, 144].

Более детальные свойства топологий, функционалов и операторов изучаются для специального класса топологических пространств — топологических линейных пространств.

§ 2. Топологические линейные пространства

Определение 1. Множество X называется *линейным пространством*, если:

А) для любой пары x, $y \in X$ определена их сумма $x+y \in X$;

Б) для любого $x \in X$ и всякого числа $\lambda \in \mathbf{R}$ ($\mathbf{R}$ — множество вещественных чисел) определено произведение $\lambda x \in X$;

В) указанные операции сложения и умножения подчинены требованиям: а) $x+y=y+x$ для любых x, $y \in X$; б) $(x+y)+z=x+(y+z)$ для любых x, y, $z \in X$; в) в X существует нулевой элемент θ, для которого $x+\theta=x$ при всяком $x \in X$; г) для каждого $x \in X$ существует противоположный элемент $(-x) \in X$ такой, что $x+(-x)=\theta$; д) $1 \cdot x=x$ при всяком $x \in X$; е) $\lambda(\mu x)=(\lambda\mu)x$ для каждого $x \in X$ и всяких λ, $\mu \in \mathbf{R}$; ж) $(\lambda+\mu)x=\lambda x+\mu x$ для любых λ, $\mu \in \mathbf{R}$ и всякого $x \in X$; з) $\lambda(x+y)=\lambda x+\lambda y$ для любых x, $y \in X$ и всякого $\lambda \in \mathbf{R}$.

Определение 2. Если X — линейное и одновременно топологическое пространство с топологией, относительно которой операции сложения $(x+y)$ и умножения на число (λx) непрерывны по обоим своим аргументам для любых x, $y \in X$, $\lambda \in \mathbf{R}$, то (X, τ) называется *топологическим линейным пространством* *).

Пусть x — фиксированный элемент из X, λ — заданное действительное число, а G ($G \subset X$) — некоторое множество. Тогда можно ввести новые множества: $x+G=\{x+y: \forall y \in G\}$, $\lambda G=\{\lambda y: \forall y \in G\}$. Специфической особенностью топологического линейного пространства X является тот факт, что каждая окрестность любой его точки x имеет вид $x+T$, где T — окрестность нулевого элемента θ. При этом, если $\mathscr{T}=\{T\}$ — база окрестностей нулевого элемента, то $x+\mathscr{T}=\{T+x\}$ — база окрестностей элемента x. Таким образом, топология в топологическом линейном пространстве определяется системой окрестностей нулевого элемента.

Определение 3. Множество G из линейного пространства называется *выпуклым*, если оно содержит вместе с любыми двумя

---

*) Подробности изложены в [90. с. 89].

своими элементами x, y и весь соединяющий их отрезок $[x, y]=\{\lambda x+(1-\lambda)y\colon \lambda\in[0, 1]\}$.

Определение 4. Хаусдорфово линейное топологическое пространство называется *локально выпуклым*, если в нем существует база окрестностей нуля, состоящая из выпуклых множеств.

Локально выпуклые пространства представляют наибольший интерес для изучения и чаще всего используются в приложениях. Их примеры будут приведены в § 3, 4.

Определение 5. Пусть X, Y — линейные пространства, а A — оператор, определенный на X и действующий из X в Y. Оператор A называется *линейным*, если $A(\lambda x+\mu y)=\lambda\cdot Ax+\mu\cdot Ay$ для любых x, $y\in X$ и всяких λ, $\mu\in\mathbf{R}$. Линейный оператор $l\colon X\to\mathbf{R}$ называется *линейным функционалом*.

Множество X^* всех линейных τ-непрерывных функционалов, определенных на топологическом линейном пространстве (X, τ), оказывается линейным пространством, если ввести операции сложения линейных функционалов и умножения их на число следующим образом: а) $(l_1+l_2)(x)=l_1(x)+l_2(x)$ для любых l_1, $l_2\in X^*$ и каждого $x\in X$; б) $(\lambda l)(x)=\lambda\cdot l(x)$ для всех $l\in X^*$ и всякого $x\in X$, $\lambda\in\mathbf{R}$.

Определение 6. Линейное пространство X^* всех линейных непрерывных функционалов, заданных на локально выпуклом пространстве X, называется *сопряженным к* X.

В локально выпуклом пространстве X кроме основной топологии (*сильной топологии*), порожденной базой $\mathcal{T}$ выпуклых окрестностей нуля, можно указать и другую локально выпуклую топологию. База окрестностей нуля $\mathcal{T}^*$ в ней определяется так. Пусть m — произвольное натуральное число, $\varepsilon>0$ — произвольное действительное число, а $l_1, \ldots, l_m$ — произвольные элементы из X^*. Тогда множество $\mathcal{T}^*$ состоит из окрестностей нуля вида $T(\varepsilon, l_1, \ldots, l_m)=\{x\in X\colon |l_i(x)|<\varepsilon,\ i=1, \ldots, m\}$.

Топология, порождаемая базой $\mathcal{T}^*$, называется *слабой*. При этом сильная топология сильнее слабой в смысле определения 1.6.

Отметим, что сходимость последовательности $\{x_n\}_{n=1}^{\infty}\subset X$ в сильной топологии (*сильная сходимость*) влечет за собой сходимость ее в слабой топологии (*слабую сходимость*) к тому же пределу. Везде далее сильная сходимость последовательности $\{x_n\}$ к элементу $x\in X$ обозначается в виде $x_n\to x$, а слабая сходимость — в виде $x_n \rightharpoonup x$.

Если X — конечномерное пространство, то сильная и слабая топологии (сходимости) совпадают.

Определение 7. Функционал J, определенный на выпуклом множестве D линейного пространства X, называется *выпуклым*, если для любой пары элементов x, $y\in D$ и для любого числа λ ($\lambda\in[0, 1]$) выполнено неравенство

$$J[\lambda x+(1-\lambda)y]\leqslant\lambda J(x)+(1-\lambda)J(y).$$

Если знак равенства в этом соотношении возможен для $x\neq y$ лишь при $\lambda=0$ и $\lambda=1$ или при $x=y$, то функционал J называется *строго выпуклым*.

Теорема 1. *Если J — строго выпуклый функционал, определенный на выпуклом множестве D линейного пространства X, и если для элементов x_1, $x_2 \in D$ выполнено равенство $J(x_1)=J(x_2)=\inf\{J(x): x\in D\}\equiv J^*$, то $x_1=x_2$.*

Подробности теории топологических линейных пространств можно почерпнуть, например, из книг [90, 93, 144, 217].

§ 3. Метрические, нормированные, банаховы и гильбертовы пространства

Часто встречающимися в приложениях примерами топологических пространств являются метрические пространства.

Определение 1. Вещественный функционал $\rho(x, y)$, определенный для каждой пары элементов x, y множества X, называется *метрикой в* X, если для него выполнены условия:

1) $\rho(x, y)\geqslant 0$ для всех x, $y\in X$, причем $\rho(x, y)=0$ тогда и только тогда, когда $x=y$;

2) метрика симметрична, т. е. $\rho(x, y)=\rho(y, x)$ для любых x, $y\in X$;

3) для любых x, y, $z\in X$ метрика удовлетворяет неравенству треугольника: $\rho(x, y)+\rho(y, z)\geqslant\rho(x, z)$.

Определение 2. Множество X с определенной на нем метрикой ρ называется *метрическим пространством*.

Определение 3. Множество $S(y, r)=\{x\in X: \rho(x, y)=r\}$ называется *сферой радиуса r с центром в точке $y\in X$*. Множество $G(y, r)=\{x\in X: \rho(x, y)<r\}$ называется *открытым шаром радиуса r с центром в точке y*.

Объединение $G(y, r)\bigcup S(y, r)=\bar{G}(y, r)$ представляет собой замыкание шара $G(y, r)$ и называется *замкнутым шаром*.

Если рассмотреть совокупность $\mathcal{T}(G)$ всех открытых шаров в X, то это множество как база топологии τ определяет в пространстве X так называемую *метрическую топологию*. Метрическое пространство X, рассматриваемое как топологическое пространство с метрической топологией, является *хаусдорфовым*. В метрических пространствах при изучении сходимости можно рассматривать только последовательности. При этом τ-сходимость последовательности $\{x_n\}$ к элементу $x\in X$ есть *сходимость по метрике*:

$$\lim_{n\to\infty}\rho(x_n, x)=0,$$

а сходимость последовательности $\{x_n\}$ к множеству D $(D\subset X)$ есть сходимость вида

$$\lim_{n\to\infty}\rho(x_n, D)=0,$$

где $\rho(x, D)=\inf\{\rho(x, y): y\in D\}$.

При изучении компактных множеств в метрическом пространстве можно рассматривать лишь τ-секвенциально компактные множества. По этой причине слово «секвенциальный» в определениях 1.11, 1.12, 1.15, 1.18 и в теореме 1.2 для метрических пространств можно опустить.

Определение 4. Множество D метрического пространства X называется *ограниченным*, если в X существует шар конечного радиуса, содержащий D.

Определение 5. Последовательность $\{x_n\}$, принадлежащая метрическому пространству X, называется *фундаментальной*, если для любого $\varepsilon>0$ найдется такой индекс N, что для всех индексов n, $m>N$ будет выполнено неравенство $\rho(x_n, x_m)<\varepsilon$.

Определение 6. Метрическое пространство X называется *полным*, если любая его фундаментальная последовательность сходится к некоторому элементу из X.

Для характеристики «близости» множеств P и Q метрического пространства X часто используют *полуотклонение* множества P от множества Q:

$$\beta(P, Q)=\sup\{\rho(x, Q):\ x\in P\},$$

а также *хаусдорфово расстояние* (*псевдометрику Хаусдорфа*) между P и Q:

$$d(P, Q)=\beta(P, Q)+\beta(Q, P).$$

Последовательность множеств $\{P_n\}$ β-сходится при $n\to\infty$ к множеству Q, если $\beta(P_n, Q)\to 0$. Последовательность $\{P_n\}$ *сходится к Q по Хаусдорфу*, если $d(P_n, Q)\to 0$.

Обратимся теперь к рассмотрению важного класса линейных топологических пространств — нормированных пространств.

Определение 7. Линейное пространство X называется *нормированным*, если для любого $x\in X$ определен вещественный функционал $\|x\|$ (или $\|x\|_X$), называемый нормой и обладающий следующими свойствами:

1) $\|x\|\geqslant 0$ для любого $x\in X$, причем $\|x\|=0$ тогда и только тогда, когда $x=\theta$;

2) для любых x, $y\in X$ выполнено неравенство треугольника: $\|x+y\|\leqslant\|x\|+\|y\|$;

3) $\|\lambda x\|=|\lambda|\cdot\|x\|$ для любого $x\in X$ и всякого числа $\lambda\in\mathbf{R}$.

Если на линейном пространстве X задана норма, то это пространство будет метрическим с метрикой $\rho(x, y)=\|x-y\|$ (*инвариантная метрика*). Метрическая топология, порождаемая инвариантной метрикой, называется *нормированной топологией*. Таким образом, нормированное пространство является метрическим и топологическим линейным пространством с нормированной топологией.

Поскольку функционал $\|x\|$ в силу определения 7 является выпуклым в X, то выпуклым множеством будет и любой шар в X, так что нормированное пространство всегда локально выпуклое. Сильная сходимость последовательности $\{x_n\}\subset X$ к элементу $x\in X$ есть *сходимость по норме*:

$$\lim_{n\to\infty}\|x_n-x\|=0,$$

а *сходимость последовательности* $\{x_n\}$ *к множеству* D $(D \subset X)$ есть сходимость вида

$$\lim_{n\to\infty} \rho(x_n, \ D) = 0,$$

где

$$\rho(x, \ D) = \inf\{\|x-y\|: \ y \in D\}.$$

Как и во всяком линейном топологическом пространстве (см. § 2), в нормированном пространстве можно ввести слабую топологию и слабую сходимость. Обратимся к классификации нормированных пространств.

Определение 8. Нормированное пространство X называется *строго нормированным*, если $\|x+y\| < \|x\| + \|y\|$ для любых линейно независимых $x, \ y \in X$.

В нормированном пространстве для любого $x \in X$ и всякого $t > 0$ можно ввести функцию

$$\mu(t, \ x) = \inf_{0<\lambda<1} \left\{ \inf \left[\frac{\lambda\|x\| + (1-\lambda)\|y\| - \|\lambda x + (1-\lambda) y\|}{\lambda(1-\lambda)} : \ y \in X, \right.\right.$$
$$\left.\left. x \neq ky, \ y \neq kx, \ k \in \mathbf{R}, \ \|x-y\| = t \right] \right\},$$

которая называется *модулем локальной выпуклости*.

Определение 9. Нормированное пространство X называется *локально равномерно выпуклым*, если $\mu(t, \ x) > 0$ для любого $x \in X$ и всякого $t > 0$.

Можно привести также модуль выпуклости

$$\delta(t) = \inf\{\mu(t, \ x): \ x \in X\}$$

и с его помощью дать

Определение 10. Нормированное пространство X называется *равномерно выпуклым*, если $\delta(t) > 0$ для всякого $t > 0$.

Очевидно, равномерно выпуклое пространство будет и локально равномерно выпуклым, а локально равномерное выпуклое пространство будет строго нормированным.

В качестве примера к определениям 7—9 рассмотрим пространство $\mathbf{R}^2$ с элементами $z = (x, \ y)$ $(x, \ y \in \mathbf{R})$ и с нормами $\|z\|_1 = |x| + |y|$, $\|z\|_2 = \sqrt{x^2 + y^2}$. Нетрудно убедиться, что пространство $\mathbf{R}^2$, снабженное первой нормой, не будет строго нормированным, в то время как при использовании второй нормы оно будет равномерно выпуклым с модулем выпуклости $\delta(t) = 1 - \sqrt{1 - t^2/4}$.

Предположим, что оператор A, действующий из нормированного пространства X в нормированное пространство Y, определен на множестве $D(A) \subset X$.

Определение 11. Оператор A называется *компактным*, если он переводит всякое замкнутое ограниченное множество D $(D \subset D(A))$ в компактное множество AD $(AD \subset Y)$.

Определение 12. Оператор A называется (*сильно*) *непрерывным* (*из* X *в* Y), если он переводит любую сходящуюся при $n \to \infty$

последовательность $\{x_n\} \subset D(A)$ $(x_n \to x \in D(A))$ в сходящуюся последовательность $\{Ax_n\} \subset Y$ $(Ax_n \to Ax \in Y)$.

Оператор A называется *слабо непрерывным*, если он переводит любую слабо сходящуюся последовательность $\{x_n\} \subset D(A)$ $(x_n \rightharpoonup x \in D(A))$ в слабо сходящуюся последовательность $\{Ax_n\}$ $(Ax_n \rightharpoonup Ax \in Y)$.

Оператор A — *усиленно непрерывный*, если он переводит всякую слабо сходящуюся последовательность $\{x_n\} \subset D(A)$ $(x_n \rightharpoonup x \in D(A))$ в сильно сходящуюся последовательность $\{Ax_n\} \subset Y$ $(Ax_n \to Ax \in Y)$.

Оператор A *деминепрерывный*, если для любой последовательности $\{x_n\} \subset D(A)$ $(x_n \to x \in D(A))$ последовательность $\{Ax_n\} \subset Y$ слабо сходится к $Ax \in Y$: $Ax_n \rightharpoonup Ax$.

Определение 13. Компактный и непрерывный оператор A: $X \to Y$ называется *вполне непрерывным*.

Определение 14. Линейный оператор A: $X \to Y$ называется *ограниченным*, если существует такая константа $M \geqslant 0$, что $\|Ax\|_Y \leqslant M\|x\|_X$ для любого $x \in X$.

Если A: $X \to Y$ — линейный ограниченный оператор, то можно определить его *норму* $\|A\| = \sup\{\|Ax\|_Y: \|x\|_X \leqslant 1\}$. Для ограниченности линейного оператора A: $X \to Y$ необходима и достаточна его непрерывность из X в Y. Если линейный ограниченный оператор компактен, то он вполне непрерывен.

Теорема 1. *Если A: $X \to Y$ — линейный вполне непрерывный оператор, то он является усиленно непрерывным из X в Y.*

В теории некорректно поставленных задач особую роль играет так называемый *оператор вложения*. Пусть X, Y — нормированные пространства, причем $X \subset Y$.

Определение 15. Оператор B: $X \to Y$, определенный на всем X, называется *оператором вложение X в Y*, если $Bx = x$ для любого $x \in X$. Если оператор вложения является компактным, то говорят, что *X компактно вложено в Y*.

Особый интерес для приложений представляет частный случай нормированного пространства — банахово пространство.

Определение 16. Полное нормированное пространство называется *банаховым*.

Пусть X — банахово пространство, а X^* — сопряженное к нему пространство, т. е. линейное пространство всех линейных непрерывных функционалов, заданных на X. Будем обозначать значение линейного непрерывного функционала $l(x) \in X^*$ в точке $x \in X$ как $<l, x>_X$ (или просто $<l, x>$). Необходимым и достаточным условием непрерывности линейного функционала $<l, x>$ является требование его ограниченности, т. е. требование существования такой константы $L \geqslant 0$, что $|<l, x>| \leqslant L\|x\|$ для любого $x \in X$. Можно ввести норму линейного непрерывного функционала в виде

$$\|l\|_{X^*} = \sup\{<l, x>_X: \|x\|_X \leqslant 1\}.$$

Тогда $|<l, x>| \leqslant \|l\|_{X^*} \cdot \|x\|$ и сопряженное пространство X^*, снабженное такой нормой, само будет банаховым. Можно поэтому рассмотреть линейное пространство X^{**}, сопряженное к X^*.

Определение 17. *Вторым сопряженным пространством к X* называется пространство X^{**} линейных непрерывных функционалов, определенных на первом сопряженном пространстве X^*.

Пространство X можно отождествить с некоторым подпространством пространства X^{**}.

Определение 18. Банахово пространство X называется *рефлексивным*, если $X^{**}=X$.

В банаховом пространстве X, как и в нормированном, вводятся понятия «сильная сходимость как сходимость по норме», а также «слабая сходимость». При этом последовательность $\{x_n\}\subset X$ слабо сходится к $x\in X$ при $n\to\infty$, если $<l, x_n>\to<l, x>$ для любого $l\in X^*$.

Определение 19. Говорят, что банахово пространство X *обладает H-свойством*, если любая последовательность $\{x_n\}\subset X$, обладающая свойствами: а) $x_n \rightharpoonup x$; б) $\|x_n\|\to\|x\|$ при $n\to\infty$, сильно сходится в X: $\|x_n-x\|\to 0$ при $n\to\infty$.

Теорема 2 [228]. *Локально равномерно выпуклое рефлексивное банахово пространство X обладает H-свойством.*

Определение 20. Нормы $\|\cdot\|$, $\|\cdot\|_1$, введенные в линейном пространстве X, называются *эквивалентными*, если существуют такие постоянные p, $q\in\mathbf{R}$, что при любом $x\in X$ выполнены неравенства. $\|x\|\leqslant p\|x\|_1$, $\|x\|_1\leqslant q\|x\|$.

Эквивалентные нормы порождают в X одну и ту же топологию. Отметим также, что все нормы в конечномерном пространстве эквивалентны.

Теорема 3 [148, 228]. *В любом рефлексивном банаховом пространстве можно ввести эквивалентную норму так, чтобы пространство стало локально равномерно выпуклым, а следовательно, строго нормированным и обладающим H-свойством.*

Рассмотрим теперь важные понятия и свойства множеств, функционалов и операторов в банаховом пространстве.

Определение 21. Множество D банахова пространства называется *замкнутым* (*слабо замкнутым*), если оно содержит любую точку, являющуюся пределом (слабым пределом) какой-либо последовательности $\{x_n\}\subset D$.

Такое определение замкнутого (слабо замкнутого) множества для банаховых пространств эквивалентно общему определению, данному в § 1. В зависимости от топологии τ (сильной или слабой) и по определениям 1.11, 1.12 можно ввести понятие сильной или слабой компактности множеств из банахова пространства. Отметим, что так определенные компактные (слабо компактные) множества будут являться замкнутыми (слабо замкнутыми).

Аналогичным образом, опираясь на определение 1.15, можно ввести понятие полунепрерывности снизу (слабой полунепрерывности снизу) функционала, определенного на множестве D банахова пространства, а также понятие непрерывности (слабой непрерывности) функционала (см. также определение 12). При этом будут справедливы следующие аналоги теоремы 1.2.

Теорема 4. *Пусть D — компактное множество банахова пространства X, а функционал $J(x)$, определенный на D, полунепрерывен* *тем снизу. Тогда число $J^*=\inf\{J(x)\colon x\in D\}$ конечно, а множество элементов, реализующих эту точную нижнюю грань, не пусто: $X_*\equiv\{x\in D\colon\ J(x)=J^*\}\neq\varnothing$. Множество X_* также компактно, и любая минимизирующая последовательность для функционала J на D сходится к X_*.*

Теорема 5. *Если D — слабо компактное множество банахова пространства X, а функционал $J(x)$, определенный на D, слабо полунепрерывен снизу на D, то число J^* конечно, множество X_* не пусто и слабо компактно. Любая минимизирующая последовательность для функционала J на D слабо сходится к X_*.*

Проверка условий слабой компактности и слабой полунепрерывности доставляет определенные трудности. Часто она осуществляется с использованием следующих утверждений [32].

Теорема 6. *Всякое ограниченное, слабо замкнутое множество рефлексивного банахова пространства слабо компактно.*

Теорема 7. *Выпуклое замкнутое множество банахова пространства слабо замкнуто.*

Для выпуклых функционалов имеется также критерий слабой полунепрерывности.

Теорема 8. *Пусть D — выпуклое множество банахова пространства. Выпуклый на D функционал $J(x)$ слабо полунепрерывен снизу на D тогда и только тогда, когда $J(x)$ полунепрерывен снизу на D.*

Заметим, что функционал $J(x)=\|x\|^{\gamma}$ $(\gamma\geqslant 1)$ слабо полунепрерывен снизу на банаховом пространстве X.

Определение 22. Функционал $J(x)$, определенный на выпуклом множестве D банахова пространства, называется *строго равномерно выпуклым на D*, если существует числовая функция $\delta(t)$, определенная при $0\leqslant t\leqslant \operatorname{diam} D\equiv\sup\{\|x-y\|\colon\ x,\ y\in D\}$, удовлетворяющая условию $\delta(+0)=0$, монотонно возрастающая в области определения и такая, что

$$J(\lambda x+(1-\lambda)y)\leqslant\lambda J(x)+(1-\lambda)J(y)-\lambda(1-\lambda)\delta(\|x-y\|)$$

для всех $x,\ y\in D$ и $\lambda\in[0,\ 1]$.

Функция $\delta(t)$ называется *модулем выпуклости* для J на D.

Строго равномерно выпуклые функционалы являются строго выпуклыми в смысле определения 2.7. Сумма выпуклого и строго равномерно выпуклого функционалов обладает свойством строгой равномерной выпуклости.

Теорема 9 [32]. *Если D — выпуклое замкнутое множество рефлексивного банахова пространства, а функционал $J(x)$ определен и строго равномерно выпуклый на D, то множество $X_*=\{x\in D\colon J(x)=J^*\}$ не пусто и состоит из единственной точки $x^*\in D$, к которой сходится всякая минимизирующая для J на D последовательность.*

Пусть теперь линейный оператор A определен на банаховом пространстве X и действует в банахово пространство Y.

Определение 23. Линейный оператор $A^*\colon\ Y^*\to X^*$ называется *сопряженным к оператору $A\colon X\to Y$*, если для любого $x\in X$ и $z\in Y^*$ выполнено равенство $\langle z,\ Ax\rangle_Y=\langle A^*z,\ x\rangle_X$.

Можно показать, что для линейного непрерывного оператора A оператор A^* также непрерывен, причем $\|A^*\|=\|A\|$.

Обратимся теперь к рассмотрению специального вида банахова пространства — гильбертова пространства.

Определение 24. Банахово пространство H называется (*вещественным*) *гильбертовым пространством*, если для любой пары элементов $x, y \in H$ определен вещественный функционал $(x, y) \equiv (x, y)_H$, который называется *скалярным произведением* этих элементов и который удовлетворяет условиям: 1) $(x, x) > 0$ при любом $x \neq \theta$; 2) $(x, y) = (y, x)$ для всех $x, y \in H$; 3) $(x+y, z) = (x, z) + (y, z)$ для любых $x, y, z \in H$; 4) $(\lambda x, y) = \lambda (x, y)$ для любых $x, y \in H$ и всякого $\lambda \in \mathbf{R}$; 5) $\|x\|^2 \equiv \|x\|_H^2 = (x, x)$ для каждого $x \in H$.

Для любой пары $x, y \in H$ выполнено *неравенство Коши — Буняковского* $|(x, y)| \leqslant \|x\| \cdot \|y\|$.

Теорема 10 (Ф. Рисс). *Всякий линейный непрерывный функционал $l(x)$, определенный на гильбертовом пространстве H, представим в виде $l(x) = (l, x)$, где элемент $l \in H$ однозначно определяется функционалом $l(x)$. При этом $\|l(x)\| = \|l\|_H$.*

Из теоремы Рисса следует, что слабая сходимость последовательности $\{x_n\} \subset H$ $(x_n \rightharpoonup x \in H)$ эквивалентна сходимости скалярных произведений $(x_n, y) \to (x, y)$ при $n \to \infty$ для любого $y \in H$. Из нее следует также, что сопряженное пространство H^* можно отождествить с H, так что гильбертово пространство заведомо рефлексивно. Оно обладает также свойством равномерной выпуклости и поэтому по теореме 2 обладает H-свойством.

Определение 25. Функционал $J(x)$, определенный на выпуклом множестве D гильбертова пространства H, называется *сильно выпуклым*, если существует такая константа $\varkappa > 0$, что

$$J(\lambda x + (1-\lambda) y) \leqslant \lambda J(x) + (1-\lambda) J(y) - \lambda (1-\lambda) \varkappa \|x - y\|^2$$

для любых $x, y \in H$ и всякого $\lambda \in [0, 1]$.

Очевидно, сильно выпуклый функционал представляет собой частный случай строго равномерно выпуклого функционала, определенного на $D \subset H$, у которого модуль выпуклости имеет вид $\delta(t) = \varkappa t^2$ $(\varkappa > 0)$. Функционал $\|x\|^2$ в гильбертовом пространстве является сильно выпуклым с константой $\varkappa = 1$.

Если A — линейный ограниченный оператор, действующий из гильбертова пространства X в гильбертово Y, то сопряженный оператор A^* действует из Y в X и определяется равенством $(y, Ax)_Y = (A^* y, x)_X$ для любых $x \in X$, $y \in Y$.

Более подробные сведения по теории метрических, нормированных, банаховых и гильбертовых пространств можно почерпнуть, например, в [79, 90, 93, 144, 211, 217].

§ 4. Примеры конкретных пространств

1. n-мерное пространство $\mathbf{R}^n$. Его элементами являются всевозможные векторы (вектор-столбцы) вида

$$x = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} \equiv (x_1, \dots, x_n)^T, \quad x_1, \dots, x_n \in \mathbf{R}.$$

Здесь и в дальнейшем символ $(...)^T$ обозначает операцию транспонирования и используется для более компактной записи. Покоординатное определение линейных операций сложения и умножения на вещественное число превращает $\mathbf{R}^n$ в линейное пространство. В $\mathbf{R}^n$ можно ввести скалярное произведение

$$(x, y)=\sum_{k=1}^{n} x_k y_k, \quad x=(x_1, ..., x_n)^T, \quad y=(y_1, ..., y_n)^T$$

и порождаемую им евклидову норму

$$\| x \| = \left\{ \sum_{k=1}^{n} x_k^2 \right\}^{1/2}.$$

Все другие нормы, которые можно ввести в $\mathbf{R}^n$, эквивалентны евклидовой. Нормированное пространство $\mathbf{R}^n$ является банаховым, а при использовании евклидовой нормы — гильбертовым.

Всякое замкнутое, ограниченное в $\mathbf{R}^n$ множество компактно в $\mathbf{R}^n$. Слабая компактность в $\mathbf{R}^n$ совпадает с сильной компактностью.

2. Пространство l_p. Состоит из всех бесконечных последовательностей действительных чисел $x=(x_1, x_2, ...)$, для которых конечна величина

$$\| x \|_{l_p} = \left\{ \sum_{k=1}^{\infty} |x_k|^p \right\}^{1/p},$$

где $p \geqslant 1$ — фиксированное число. Линейные операции определяются покоординатно. При $p>1$ пространство l_p рефлексивное, строго нормированное банахово пространство, сопряженным к которому будет пространство l_q $(q \equiv p/(p-1))$. При $p=2$ пространство l_2, снабженное скалярным произведением

$$(x, y)=\sum_{k=1}^{\infty} x_k y_k, \quad x, y \in l_2,$$

является гильбертовым.

Теорема 1 (критерий компактности в l_p). *Для компактности замкнутого множества $K \subset l_p$ необходимо и достаточно, чтобы множество K было ограниченным и чтобы для любого $\varepsilon>0$ существовал такой номер $N(\varepsilon)$, что*

$$\sum_{k=n+1}^{\infty} |x_k|^p < \varepsilon^p$$

для всякого $n \geqslant N(\varepsilon)$ и для любого $x=(x_1, x_2, ...) \in K$.

Теорема 2. *Для того чтобы последовательность $\{x^{(n)}\}$ элементов $x^{(n)}=(\xi_1^{(n)}, \xi_2^{(n)}, ...) \in l_p$ $(p>1)$ слабо сходилась к $x_0=(\xi_1^{(0)}, \xi_2^{(0)}, ...)$, необходимо и достаточно, чтобы:* а) *последовательность норм $\{\| x^{(n)} \|\}$ была ограничена*; б) $\xi_k^{(n)} \to \xi_k^{(0)}$ *при $n \to \infty$ для всех $k=1, 2, ...$*

3. Пространство $C[a, b]$. Состоит из вещественных функций $x(t)$, непрерывных на отрезке $[a, b]$ и имеющих норму $\| x \|_C = \max\{|x(t)|: t \in [a, b]\}$. При этом сложение функций и их умножение на число определяется обычным образом. Пространство $C[a, b]$ банахово, однако оно не является рефлексивным и строго нормированным. Сильная сходимость последовательности $\{x_n\} \subset C[a, b]$ к x_0 означает равномерную сходимость функциональной последовательности $\{x_n(t)\}$ к функции $x_0(t) \in C[a, b]$.

Пусть имеется некоторое множество функций D $(D \subset C[a, b])$. Функции $x(t)$ из множества D называются *равномерно ограниченными*, если существует такая константа $M \geqslant 0$, что $\|x\|_C \leqslant M$ для любого $x \in D$. Функции $x(t)$ из D называются *равностепенно непрерывными*, если для любого $\varepsilon > 0$ существует число $\delta(\varepsilon) > 0$ такое, что для любых t_1, $t_2 \in [a, b]$, удовлетворяющих неравенству $|t_1 - t_2| < \delta$, и для всякой функции $x(t) \in D$ выполнено неравенство $|x(t_1) - x(t_2)| < \varepsilon$.

Теорема 3 (критерий компактности в $C[a, b]$). *Для того чтобы замкнутое множество D $(D \subset C[a, b])$ было компактным в $C[a, b]$ необходимо и достаточно, чтобы функции $x(t) \in D$ были равномерно ограничены и равностепенно непрерывны.*

Эту теорему часто называют *теоремой Арцела*. Слабая сходимость в $C[a, b]$ характеризуется следующим образом.

Теорема 4. *Для того чтобы последовательность $\{x_n\} \subset C[a, b]$ слабо сходилась к $x_0 \in C[a, b]$, необходимо и достаточно, чтобы* а) *функции из множества $\{x_n(t)\}$ были равномерно ограниченными* б) $x_n(t) \to x_0(t)$ *при* $n \to \infty$ *для любого* $t \in [a, b]$.

4. Пространства $C^m[a, b]$. Пусть $m \geqslant 1$ — фиксированное число Тогда нормированное пространство $C^m[a, b]$ состоит из m раз непрерывно дифференцируемых функций $x(t)$ с нормой

$$\|x\|_{C^m} = \|x\|_C + \sum_{k=1}^{m} \|x^{(k)}\|_C.$$

Формально считают $C^0[a, b] = C[a, b]$. Пространства $C^m[a, b]$ банаховы. Они не являются строго нормированными и рефлексивными.

5. Пространства Гёльдера $H^\mu[a, b]$. Пространство $H^\mu[a, b]$ состоит из всех функций $x(t)$, заданных на $[a, b]$ и удовлетворяющих там условию Гёльдера (или Липшица) с показателем μ $(0 < \mu \leqslant 1)$, т. е. из функций, для которых конечна величина

$$\sup\{|x(t_1) - x(t_2)|/|t_1 - t_2|^\mu\colon\ t_1,\ t_2 \in [a, b],\ t_1 \neq t_2\}.$$

Норма в $H^\mu[a, b]$ определяется в следующем виде:

$$\|x\|_{H^\mu} = \|x\|_C + \sup\{|x(t_1) - x(t_2)|/|t_1 - t_2|^\mu\colon\ t_1,\ t_2 \in [a, b],\ t_1 \neq t_2\}.$$

При этом $H^\mu[a, b]$ становится банаховым пространством. Если $0 < \nu < \mu \leqslant 1$, то справедливы вложения

$$C^1[a, b] \subset H^\mu[a, b] \subset H^\nu[a, b] \subset C[a, b].$$

Последнее вложение компактное. Пространства $H^\mu[a, b]$ не являются рефлексивными и строго нормированными.

6. Пространства $L_p[a, b]$ (пространства Лебега). Множество всех измеримых по Лебегу функций (определение и теорию измеримости по Лебегу см., например, в [93, 144, 211]) $x(t)$ $(t \in [a, b])$, для которых конечен интеграл

$$\int_a^b |x(t)|^p \, dt \quad (p \geqslant 1),$$

называется *пространством функций, суммируемых с p-й степенью на* $[a, b]$, или, короче, $L_p[a, b]$. Пространство $L_p[a, b]$ становится банаховым, если ввести норму

$$\|x\|_{L_p} = \left\{ \int_a^b |x(t)|^p \, dt \right\}^{1/p}.$$

При этом элементы из $L_p[a, b]$, отличающиеся на множестве меры нуль (см., например, [93]), отождествляются. При $p > 1$ пространство $L_p[a, b]$ рефлексивное и равномерно выпуклое. Таким образом, оно является также строго нормированным и обладает H-свойством. Сопряженным к $L_p[a, b]$ при $p > 1$ будет пространство $L_q[a, b]$, где $q = p/(p-1)$. Пространство $L_2[a, b]$ со скалярным произведением

$$(x, y)_{L_2} = \int_a^b x(t) y(t) \, dt, \quad x, y \in L_2[a, b],$$

является гильбертовым.

Пространство $L_1[a, b]$ не является строго нормированным. Справедливы вложения $L_p[a, b] \subset L_1[a, b]$, $C[a, b] \subset L_p[a, b]$ при $p \geqslant 1$.

Теорема 5 (критерий компактности в $L_p[a, b]$). *Для того чтобы замкнутое множество функций* $D = \{x(t)\} \subset L_p$ *было компактным, необходимо и достаточно, чтобы функции из* D *обладали свойствами*:

а) $\|x\|_{L_p} \leqslant M = \text{const}$ *для любого* $x \in D$ (*равномерная ограниченность норм*);

б) *для любого* $\varepsilon > 0$ *найдется такое число* $\delta(\varepsilon) > 0$, *что для всех функций* $x \in D$ *и при всех* h $(0 < h < \delta(\varepsilon))$ *выполнено неравенство*

$$\int_a^b |x(t+h) - x(t)|^p \, dt < \varepsilon^p$$

(*равностепенная непрерывность в среднем*).

Эта теорема называется *критерием Рисса*. Имеются другие критерии компактности в L_p (см., например, [144, 217]).

Слабую сходимость в $L_p[a, b]$ можно охарактеризовать следующей теоремой.

Теорема 6. *Для того чтобы последовательность* $\{x_n(t)\} \subset L_p[a, b]$ $(p > 1)$ *слабо сходилась к элементу* $x_0(t) \in L_p[a, b]$, *необходимо и достаточно, чтобы*: а) *последовательность* $\{\|x_n\|\}$ *была ограничена*; б) *для любого* $s \in [a, b]$ *имела место сходимость* $\int_a^s x_n(t) \, dt \to \int_a^s x_0(t) \, dt$ *при* $n \to \infty$.

Функции из пространства $L_1[a, b]$, а значит, и функции из $L_p[a, b]$ обладают рядом важных свойств.

Теорема 7. *Если функция* $\varphi(t)$ *принадлежит* $L_1[a, b]$ *и почти всюду на* $[a, b]$ *(т. е. в каждой точке отрезка* $[a, b]$, *за исключением множества меры* 0) *выполнено неравенство* $|x(t)|\leqslant\varphi(t)$, *то* $x\in L_1[a, b]$. *В частности, если* $|x(t)|\leqslant M=\text{const}$ *почти всюду, то* $x\in L_1[a, b]$.

Теорема 8 (А. Лебег). *Если последовательность функций* $\{x_n(t)\}$ *сходится почти всюду при* $n\to\infty$ *к* $x(t)$ *и для каждого* n *почти всюду на* $[a, b]$ *выполнено неравенство* $|x_n(t)|\leqslant\varphi(t)$, *где* $\varphi(t)\in L_1[a, b]$, *то предельная функция* $x(t)$ *принадлежит пространству* $L_1[a, b]$. *При этом*

$$\lim_{n\to\infty}\int_a^b x_n(t)\,dt=\int_a^b x(t)\,dt,\quad \lim_{n\to\infty}\|x_n\|_{L_1}=\|x\|_{L_1}.$$

Следствие. *Если* $|x_n(t)|\leqslant M$ *и* $x_n(t)\to x(t)$ *почти всюду на* $[a, b]$, *то* $x\in L_1[a, b]$ *и справедливы приведенные в теореме Лебега предельные соотношения.*

Приведем также данные о связи различных типов сходимости последовательности $\{x_n(t)\}\subset L_2[a, b]$ к функции $x(t)\in L_2[a, b]$. Так, из равномерной сходимости $x_n(t)\rightrightarrows x(t)$ следуют сходимости $x_n(t)\overset{L_2}{\to}x(t)$, $x_n(t)\to x(t)$ (почти) всюду на $[a, b]$. Из сходимости $x_n(t)\overset{L_2}{\to}x(t)$ вытекает сходимость $x_n(t)\overset{L_1}{\to}x(t)$, а из последней — сходимость по мере [93, с. 285]. Из сходимости почти всюду также следует сходимость по мере. Подробно этот вопрос изложен в [93, с. 378].

В пространствах $L_p[a, b]$ функционал $J(x)=\|x\|^{\gamma}_{L_p}$ строго равномерно выпукл на всем пространстве при любом γ $(\gamma\geqslant p\geqslant 2)$ с модулем выпуклости $\delta(t)=(1/2)^{\gamma-1}t^{\gamma}$. При $1<p<2$ этот функционал строго равномерно выпукл для всякого $\gamma>1$ на любом выпуклом ограниченном множестве из $L_p[a, b]$.

Пространства L_p можно ввести также и для случая функций нескольких переменных. Пусть $T\subset\mathbf{R}^n$ — замкнутое ограниченное множество. Множество всех измеримых по Лебегу функций $x(t)=x(t_1, ..., t_n)$ $(t\equiv(t_1, ..., t_n)\in T)$, для которых конечен интеграл

$$\int_T|x(t_1, ..., t_n)|^p\,dt_1\ldots dt_n\equiv\int_T|x(t)|^p\,dt,$$

называется *пространством* $L_p(T)$. Норма в $L_p(T)$ вводится равенством

$$\|x\|_{L_p(T)}=\left\{\int_T|x(t)|^p\,dt\right\}^{1/p}.$$

В случае $p=2$ можно ввести также скалярное произведение

$$(x, y)_{L_2(T)}=\int_T x(t)\,y(t)\,dt.$$

Пространства $L_p(T)$ являются банаховыми, а пространство $L_2(T)$ гильбертовым. Свойства этих пространств аналогичны свойствам пространств $L_p[a, b]$, $L_2[a, b]$.

7. Пространства W_p^l (пространства Соболева). Пусть $T\subset\mathbf{R}^n$ — замкнутое ограниченное множество, а $x(t)$, $y(t)$ $(t\equiv(t_1, ..., t_n))$ — функции, определенные на T. Известно, что если обе эти функции являются бесконечно дифференцируемыми в T, причем вторая и ее производные обращаются в нуль на границе ∂T, то справедливо равенство

$$\int_T \left\{ x(t) \frac{\partial^k y}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} + (-1)^{k+1} y(t) \frac{\partial^k x}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} \right\} dt = 0, \quad k = k_1 + \dots + k_n,$$

которое получается с помощью k-кратного интегрирования по частям.

Определение 1. Интегрируемая по Лебегу в T функция $z_{k_1 \dots k_n}(t)$ называется *обобщенной производной вида* $\partial^k / \partial t_1^{k_1} \dots \partial t_n^{k_n}$ *функции* $x(t)$, если для любой бесконечно дифференцируемой функции $y(t)$, обращающейся в нуль вместе со всеми своими производными в некоторой окрестности границы области T, имеет место равенство

$$\int_T \left\{ x(t) \frac{\partial^k y}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} + (-1)^{k+1} y(t) z_{k_1 \dots k_n}(t) \right\} dt = 0.$$

Множество всех функций из $L_p(T)$ $(p \geqslant 1)$, имеющих все обобщенные производные до l-го порядка включительно, которые также принадлежат $L_p(T)$, называется *пространством* $W_p^l(T)$. Норма в $W_p^l(T)$ вводится согласно равенству

$$\| x \|_{W_p^l(T)} = \left\{ \int_T | x(t) |^p dt + \sum_{\substack{1 \leqslant k \leqslant l \\ k = k_1 + \dots + k_n}} \int_T \left| \frac{\partial^k x}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} \right|^p dt \right\}^{1/p}.$$

Определенное таким образом нормированное пространство $W_p^l(T)$ является рефлексивным, равномерно выпуклым банаховым пространством при $p > 1$ и для любого $l \geqslant 1$. Поэтому оно обладает H-свойством.

Особый интерес представляют собой пространства $W_2^l(T)$, которые будут гильбертовыми, если определить в них скалярное произведение

$$(x, y)_{W_2^l(T)} = (x, y)_{L_2(T)} + \sum_{\substack{1 \leqslant k \leqslant l \\ k = k_1 + \dots + k_n}} \int_T \frac{\partial^k x}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} \frac{\partial^k y}{\partial t_1^{k_1} \dots \partial t_n^{k_n}} dt.$$

Если ввести пространство $C(T)$ непрерывных на T функций $x(t)$ с нормой $\| x \|_{C(T)} = \max \{ | x(t) |: \ t \in T \}$, то справедлива следующая теорема вложения.

Теорема 9 [178]. *При условии $2l > n$ пространство $W_2^l(T)$ компактно вложено в $C(T)$ и $\| x \|_{C(T)} \leqslant k_0 \| x \|_{W_2^l(T)}$ для любого $x \in W_2^l(T)$* $(k_0 = \text{const} > 0)$.

Рассмотрим специально случай $n = 1$, когда $T = [a, b]$. В этом случае можно нормировать пространства $W_p^l[a, b]$ следующим образом:

$$\| x \|_{W_p^l[a, b]} = \left\{ \int_a^b q(t) | x(t) |^p dt + \sum_{k=1}^{l} \int_a^b r_k(t) | x^{(k)}(t) |^p dt \right\}^{1/p},$$

где функции $q(t)$, $r_k(t)$ $(k=1, ..., l)$ непрерывны и положительны на $[a, b]$. Эквивалентная норма имеет вид

$$\| x \|_{W_p^l[a, b]} = \{ \| x \|^p_{L_p[a, b]} + \| x^{(l)} \|^p_{L_p[a, b]} \}^{1/p}.$$

Скалярное произведение в гильбертовом пространстве $W_2^1[a, b]$ вводится так:

$$(x, y)_{W_2^1[a, b]} = (x, y)_{L_2[a, b]} + (x', y')_{L_2[a, b]} = \int_a^b \{x(t)y(t) + x'(t)y'(t)\}\, dt.$$

Согласно теореме 9, гильбертово пространство $W_2^1[a, b]$ компактно вложено в $C[a, b]$, причем $\| x \|_{C[a, b]} \leqslant k_0 \| x \|_{W_2^1[a, b]}$, где $k_0 = \max(\sqrt{b-a}, (b-a)^{-1/2})$. Вложение $W_2^1[a, b] \subset C[a, b]$ понимается так, что любой элемент $x(t) \in W_2^1$, являясь функцией из $L_2[a, b]$, почти всюду совпадает с некоторой непрерывной функцией и поэтому с ней отождествляется.

Теорема 10 [211]. *Любая функция $x(t)$ из $W_2^1[a, b]$ абсолютно непрерывна, т. е. почти всюду на $[a, b]$ имеет конечную обычную производную $x'(t)$, которая совпадает с обобщенной производной, причем*

$$x(t) = x(a) + \int_a^t x'(\tau)\, d\tau$$

для любого $t \in [a, b]$.

Отметим также следующую теорему вложения пространств Соболева при $n=1$.

Теорема 11. 1) *Если l, p, q, $m \geqslant 1$, $1 > (l-m)p$, $q < p/[1-(l-m)p]$, то вложение $W_p^l[a, b] \subset W_q^m[a, b]$ компактно.*

2) *При условиях $p \geqslant 1$, $lp > 1$, $0 < \mu < 1$, $\mu < (lp-1)/p$ вложение $W_p^l[a, b] \subset H^\mu[a, b]$ компактно.*

В частности, $W_2^1[a, b]$ компактно вложено в $H^\mu[a,b]$ при $0 < \mu < 1/2$.

Указанные теоремы о компактных вложениях часто используются в теории некорректных задач.

8. Пространство $V[a, b]$. Пусть на $[a, b]$ задана конечная функция $x(t)$. Зададим произвольное разбиение отрезка $[a, b]$ на части точками $a = t_1 < ... < t_N = b$. Обозначим это разбиение символом Π_N $(\Pi_N = \{t_1, ..., t_N\})$. Составим сумму

$$\omega(x, \Pi_N) = \sum_{k=1}^{N-1} |x(t_{k+1}) - x(t_k)|.$$

Рассмотрим теперь множество Π всех возможных разбиений отрезка $[a, b]$ и множество всех соответствующих этим разбиениям сумм вида $\omega(x, \Pi_N)$.

Определение 2. *Вариацией (полным изменением) функции $x(t)$ на отрезке $[a, b]$* называется величина

$$\bigvee_a^b (x) = \sup\{\omega(x, \Pi_N):\ \Pi_N \in \Pi\}.$$

Если для данной функции $x(t)$ вариация конечна, то эта функция называется *функцией с ограниченной вариацией (конечным изменением).*

Пространство функций $V[a, b]$ состоит из всех функций, имеющих на $[a, b]$ ограниченную вариацию. Нормируя $V[a, b]$ следующим образом:

$$\|x\|_{V[a,b]}=|x(a)|+\bigvee_a^b(x)\quad \forall x\in V[a, b],$$

получаем банахово пространство. Оно не является строго нормированным и рефлексивным. Функции из $V[a, b]$ обладают рядом важных свойств [93, 168].

Теорема 12. *Для включения $x(t)\in V[a, b]$ необходимо и достаточно, чтобы функция $x(t)$ представлялась в виде разности двух монотонно не убывающих ограниченных функций.*

Поэтому любая функция из $V[a, b]$ может иметь не более чем счетное множество точек разрыва первого рода.

Теорема 13. *Если $x(t)\in V[a, b]$, то почти всюду на $[a; b]$ существует конечная производная $x'(t)$, причем $x'(t)\in L_1[a, b]$. Если $x(t)$ всюду на $[a, b]$ имеет конечную производную $x'(t)$, причем $x'(t)\in L_1[a, b]$, то $x(t)\in V[a, b]$.*

Теорема 14. *Если $x(t)\in V[a, b]$ и $x(t)$ непрерывна в некоторой точке $t_0\in[a, b]$, то в этой же точке непрерывна функция*

$$q(t)=\bigvee_a^t(x).$$

Вариация монотонной на $[a, b]$ функции $x(t)$ равна величине $|x(b)-x(a)|$.

Теорема 15 (Хелли о компактности). *Из всякого бесконечного множества функций $\mathcal{F}$ ($\mathcal{F}\subset V[a, b]$), ограниченного в $V[a, b]$, можно выбрать функциональную последовательность, которая сходится в каждой точке отрезка $[a, b]$ к функции из $V[a, b]$.*

Предположим, что функция $f(t)$ непрерывна на $[a, b]$, а $x(t)\in V[a, b]$. Тогда определен *интеграл Стилтьеса* [93, 168]

$$\int_a^b f(t)\,dx(t).$$

Теорема 16 (Ф. Рисс, см. [93, с. 362]). *При сделанных предположениях о $f(t)$ и $x(t)$ справедлива оценка*

$$\left|\int_a^b f(t)\,dx(t)\right|\leqslant\|f\|_{C[a,b]}\cdot\|x\|_{V[a,b]}.$$

Интеграл Стилтьеса не зависит от значений, принимаемых функцией $x(t)$ в ее точках разрыва. Для интеграла Стилтьеса справедлива следующая теорема о предельном переходе под знаком интеграла.

Теорема 17 (Хелли о предельном переходе). *Пусть $f(t)\in C[a, b]$, а последовательность функций $\{x_n(t)\}\subset V[a, b]$ ограничена по норме и сходится в каждой точке отрезка $[a, b]$ к функции $x(t)$. Тогда*

$$\lim_{n\to\infty}\int_a^b f(t)\,dx_n(t)=\int_a^b f(t)\,dx(t).$$

При этом по теореме 15 $x(t)\in V[a, b]$.

В заключение отметим очевидные вложения $C^1[a, b] \subset V[a, b] \subset L_p[a, b]$ $(p \geqslant 1)$.

Более детально функциональные пространства, приведенные в данном параграфе, описаны, например, в книгах [93, 144, 168, 178, 211, 217].

§ 5. Элементы дифференциального исчисления в банаховом пространстве

Пусть X — банахово пространство и функционал $J(x)$ определен на некотором открытом множестве $\tilde{D}$ $(\tilde{D} \subset X)$, содержащем множество D.

Определение 1. *Функционал $J(x)$ дифференцируем в точке $x \in D$*, если существует элемент $J'(x) \in X^*$ такой, что для любого h $(h \in X, x+h \in \tilde{D})$ справедливо равенство

$$\Delta J(x) = J(x+h) - J(x) = \langle J'(x), h \rangle_X + w(x, h).$$

При этом функционал $w(x, h)$, определенный для указанных h, удовлетворяет условию $w(x, h)/\|h\|_X \to 0$ при $\|h\|_X \to 0$, т. е. $w(x, h) = o(\|h\|)$.

Если функционал J дифференцируем в каждой точке множества D, то он называется *дифференцируемым на D*.

Величина $dJ(x) = \langle J'(x), h \rangle$ называется (*первым*) *дифференциалом* функционала $J(x)$ в точке x, а элемент $J'(x)$ — *градиентом* (*первой производной*) этого функционала в точке x. Если градиент существует в данной точке, то он определяется однозначно.

Предположим, что функционал $J(x)$ дифференцируем в точке x. Тогда числовая функция $\varphi(\lambda) = J(x+\lambda h) - J(x)$, определенная при $\lambda \in [0, 1]$, дифференцируема и

$$dJ(x) = \langle J'(x), h \rangle = \varphi'(0) = \lim_{\lambda \to +0} \frac{J(x+\lambda h) - J(x)}{\lambda} \quad \forall h \in X.$$

Определение 2. *Функционал $J(x)$* называется *непрерывно дифференцируемым на множестве D*, если он дифференцируем на D и $\|J'(x+h) - J'(x)\|_{X^*} \to 0$ при $\|h\|_X \to 0$ для всех x, $x+h \in D$.

Для приложений важную роль играют следующие примеры.

Пример 1. Пусть X — гильбертово пространство и $J(x) = \|x\|^2$. Тогда этот функционал непрерывно дифференцируем на X и $J'(x) = 2x$. Если же $J(x) = \|x\|$, то функционал дифференцируем всюду, кроме нуля, и $J'(x) = x/\|x\|$. В нулевой точке функционал $\|x\|$ не является дифференцируемым.

Пример 2. Пусть A: $X \to Y$ — линейный непрерывный оператор, действующий из банахова пространства X в гильбертово пространство Y. Если $J(x) = \|Ax - y\|_Y^2$ $(y \in Y)$ — определенный на X функционал, то он непрерывно дифференцируем на X и $J'(x) = 2A^*(Ax - y)$.

Пример 3. Пусть функция $K(x, s, z)$ определена, непрерывна и имеет непрерывную частную производную $K'_z(x, s, z)$ в области $\Pi = [c, d] \times [a, b] \times \mathbf{R}$. Тогда для любой функции $z(s) \in C[a, b]$

определен интегральный оператор

$$A[x,\ z(s)]=\int_a^b K(x,\ s,\ z(s))ds, \quad x\in[c,\ d],$$

действующий в пространство $C[c, d]$. Зададим на $C[a, b]$ функционал $J(z)=\|A[x, z(s)]-u(x)\|^2_{L_2[c,d]}$, где $u(x)\in L_2[c, d]$. Этот функционал будет непрерывно дифференцируемым на $C[a, b]$, причем

$$J'(z)=2\int_c^d K'_z(x,\ s,\ z(s))\left[\int_a^b K(x,\ t,\ z(t))dt-u(x)\right]dx.$$

Определение 3. *Функционал $J(x)$ дважды дифференцируем в точке $x\in D$*, если приращение $\Delta J(x)=J(x+h)-J(x)$ для всех $h\in X$, $x+h\in\tilde{D}$ представимо в виде

$$\Delta J(x)=\langle J'(x),\ h\rangle+\frac{1}{2}\langle J''(x)h,\ h\rangle+\gamma(h,\ x).$$

Здесь $J'(x)$ — градиент функционала $J(x)$ в точке x, а $J''(x)$ — линейный ограниченный оператор, действующий из X в X^*. Функционал $\gamma(h, x)$ определен для указанных h и удовлетворяет условию $\gamma(h, x)/\|h\|^2\to 0$ при $\|h\|\to 0$.

Квадратичная форма $d^2J(x)=\langle J''(x)h,\ h\rangle$ называется *вторым дифференциалом* функционала $J(x)$ в точке x, а оператор $J''(x)$ — *второй производной* функционала $J(x)$ в точке x.

Определение 4. Функционал $J(x)$ называется *дважды непрерывно дифференцируемым на множестве D*, если он дважды дифференцируем в каждой точке этого множества и $\|J''(x+h)-J''(x)\|\to 0$ при $\|h\|\to 0$ для всех $x\in D$ и $x+h\in D$.

В частности, функционалы $F(x)=\|x\|^2$, $J(x)=\|Ax-y\|^2$ из примеров 1, 2 будут дважды непрерывно дифференцируемыми на X, причем их вторые производные имеют соответственно вид $J''(x)=2E$, $J''(x)=2A^*A$.

Здесь E — единичный оператор, действующий из X в X.

Введенный в определениях 1—4 понятия дифференцируемости позволяют сформулировать специальные критерии выпуклости функционалов [32].

Теорема 1. *Если D — выпуклое множество банахова пространства X, а $J(x)$ — непрерывно дифференцируемый на D функционал, то для выпуклости его на D необходимо и достаточно выполнения для любых x, $y\in D$ одного из следующих неравенств*:

$$J(x)\geqslant J(y)+\langle J'(y),\ x-y\rangle,\quad \langle J'(x)-J'(y),\ x-y\rangle\geqslant 0.$$

Если, кроме того, множество D имеет внутренние точки и функционал $J(x)$ дважды непрерывно дифференцируем на D, то для его выпуклости на D необходимо и достаточно, чтобы выполнялось неравенство $\langle J''(x)h,\ h\rangle\geqslant 0$ для любых $x\in D$ и $h\in X$.

Теорема 2. *Если D — выпуклое множество гильбертова пространства X и функционал $J(x)$ непрерывно дифференцируем на D, то для сильной выпуклости $J(x)$ на D необходимо и достаточно, чтобы* *существовала такая константа* $\varkappa>0$, *что для любых* x, $y\in D$ *выполнено неравенство*

$$(J'(x)-J'(y),\ x-y)\geqslant 2\varkappa\|x-y\|^2.$$

Если множество D *имеет внутренние точки и функционал* $J(x)$ *дважды непрерывно дифференцируем на* D, *то для его сильной выпуклости на* D *необходимо и достаточно существование константы* $\varkappa>0$ *такой, что для любых* $x\in D$ *и* $h\in X$ *выполнено неравенство* $(J''(x)h,\ h)\geqslant 2\varkappa\|h\|^2$.

Из теоремы 2, в частности, ясно, что функционал $J(x)=\|x\|^2$, фигурирующий в примере 1, будет сильно выпуклым в гильбертовом пространстве X, причем $\varkappa=1$.

Приведем также необходимые и достаточные условия минимума дифференцируемого функционала на выпуклом множестве.

Теорема 3. *Пусть* D — *выпуклое множество банахова пространства* X, *а* $J(x)$ — *непрерывно дифференцируемый на* D *функционал. Для того чтобы точка* $x^*\in D$ *удовлетворяла равенству* $J(x^*)=\inf\{J(x):\ x\in D\}\equiv J^*$, *необходимо выполнение неравенства*

$$\langle J'(x^*),\ x-x^*\rangle\geqslant 0 \tag{1}$$

для всех $x\in D$. *В случае, когда* x^* — *внутренняя точка множества* D, *условие* (1) *эквивалентно равенству* $J'(x^*)=0$. *Если функционал* $J(x)$, *кроме того, выпуклый на* D, *то условие* (1) *является также и достаточным для того, чтобы элемент* $x^*\in D$ *удовлетворял равенству* $J(x^*)=J^*$.

Пример 4. Предположим, что линейный оператор A действует из гильбертова пространства X в гильбертово пространство Y. Рассмотрим функционал

$$M^\alpha[x]=\alpha\|x\|_X^2+\|Ax-y\|_Y^2,\quad y\in Y,\ \alpha>0,$$

определенный при $x\in X$. Как отмечено выше, функционал $\|x\|_X^2$ непрерывно дифференцируемый и сильно выпуклый на X, а функционал $\|Ax-y\|_Y^2$ непрерывно дифференцируемый и выпуклый на X. Поэтому по теореме 3 необходимым и достаточным условием того, что элемент $x^\alpha\in X$ реализует точную нижнюю грань функционала $M^\alpha[x]$, является равенство $(M^\alpha)'[x]=2[(\alpha E+A^*A)x^\alpha-A^*y]=0$, или

$$(\alpha E+A^*A)x^\alpha=A^*y. \tag{2}$$

Поскольку функционал $M^\alpha[x]$ сильно выпуклый, то по теореме 3.9 элемент x^α существует и единствен. Уравнение (2) удобно для практического отыскания x^α.

Подробное изложение дифференциального исчисления в банаховых пространствах и детали его приложения к исследованию функционалов можно найти, например, в книгах [31, 32, 90, 93, 144].

§ 6. Некоторые сведения из линейной алгебры

В этом параграфе будут приведены некоторые сведения из линейной алгебры, которые широко используются в гл. 5. Пусть $A=[a_{ij}]_{m\times n}$ — действительная матрица с элементами a_{ij} ($i=1,\ \dots,\ m$; $j=1,\ \dots,\ n$) размера $m\times n$ ($\dim A=m\times n$), а $z=(z_1,\ \dots,\ z_n)^T$,

$u=(u_1, \ldots, u_m)^T$ — вектор-столбцы в соответственно пространствах $\mathbf{R}^n$, $\mathbf{R}^m$, снабженных евклидовой нормой. При изучении системы линейных алгебраических уравнений (СЛАУ) вида

$$Az=u \tag{1}$$

относительно неизвестных $(z_1, \ldots, z_n)$ возможна ситуация, когда СЛАУ (1) не имеет решений (в классическом смысле). При этом, однако, всегда существуют обобщенные решения СЛАУ (1), называемые *псевдорешениями*. Вектор $\tilde{z}\in\mathbf{R}^n$ называется псевдорешением СЛАУ (1), если для него выполнено равенство

$$\|A\tilde{z}-u\|^2=\inf\{\|Az-u\|^2:\ z\in\mathbf{R}^n\}. \tag{2}$$

Нахождение псевдорешений из условия (2) называется *методом наименьших квадратов* [231, 237] и эквивалентно решению нормальной системы уравнений

$$A^TAz=A^Tu.$$

Если система (1) разрешима, то ее псевдорешения совпадают с обычными решениями. Псевдорешение СЛАУ (1) может и не быть единственным.

Нормальным псевдорешением системы (1) называется псевдорешение с наименьшей нормой. Нормальное псевдорешение существует и единственно для любой СЛАУ (1). Если система (1) разрешима, то ее нормальное псевдорешение называется *нормальным решением*.

Будем обозначать символом $\mathrm{Ker}\,A$ множество всех решений однородной системы $Az=0$, а символом $\mathrm{Im}\,A$ множество всех векторов вида $\{Az\}$. Очевидно, $\mathrm{Ker}\,A\subset\mathbf{R}^n$, $\mathrm{Im}\,A\subset\mathbf{R}^m$. Справедливы также соотношения

$$\mathrm{Ker}\,A=(\mathrm{Im}\,A^T)^\perp,\qquad \mathrm{Ker}\,A^T=(\mathrm{Im}\,A)^\perp,\quad \mathrm{Ker}(A^TA)=\mathrm{Ker}\,A,$$

$$\mathrm{Im}(A^TA)=\mathrm{Im}\,A^T,\qquad \mathrm{Ker}(AA^T)=\mathrm{Ker}\,A^T,\qquad \mathrm{Im}(AA^T)=\mathrm{Im}\,A.$$

Здесь символом $D^\perp$ обозначено ортогональное дополнение множества D из пространства $\mathbf{R}^n(\mathbf{R}^m)$, т. е. совокупность всех таких $y\in\mathbf{R}^n(\mathbf{R}^m)$, что $(x, y)=0$ для любого $x\in D$.

Теорема 1. *Для того чтобы вектор* $\tilde{z}\in\mathbf{R}^n$ *был псевдорешением СЛАУ* (1), *необходимо и достаточно, чтобы вектор* $A\tilde{z}-u$ *был ортогонален множеству* $\mathrm{Im}\,A$.

Для того чтобы псевдорешение $\bar{z}$ *СЛАУ* (1) *было нормальным, необходимо и достаточно, чтобы существовал такой элемент* $\bar{v}\in\mathbf{R}^m$, *что* $\bar{z}=A^T\bar{v}$ (*или, что то же самое, необходимо и достаточно, чтобы* $\bar{z}\perp\mathrm{Ker}\,A$).

Зафиксируем матрицу A и рассмотрим систему (1) для различных $u\in\mathbf{R}^m$. Если каждому вектору $u\in\mathbf{R}^m$ поставить в соответствие нормальное псевдорешение $\bar{z}$ СЛАУ (1), то это соответствие представляет собой однозначное линейное отображение пространства $\mathbf{R}^m$ в пространство $\mathbf{R}^n$, которое осуществляется некоторым оператором. Матрица этого оператора, имеющая размер $n\times m$, обозначается как A^+ и называется *псевдообратной матрицей* для A. Если $\dim A=n\times n$, $\det A\neq 0$, то $A^+=A^{-1}$, т. е. псевдообратная матрица совпадает с существующей в этом случае обычной обратной матрицей. Поэтому псевдообратную матрицу называют еще *обобщенной обратной*. Если определитель квадратной матрицы $A^T A$ не равен нулю, то $A^+ = (A^T A)^{-1} A^T$.

Для любой матрицы A существует единственная псевдообратная матрица A^+. В силу определения нормальное псевдорешение $\bar{z}$ системы (1) с правой частью $u \in \mathbf{R}^m$ вычисляется по формуле $\bar{z} = A^+ u$.

Операция нахождения матрицы, псевдообратной для A, называется *псевдообращением*. Приведем основные свойства операции псевдообращения.

Теорема 2 [46]. 1) $\operatorname{Ker} A^+ = \operatorname{Ker} A^T$, $\operatorname{Im} A^+ = \operatorname{Im} A^T$; 2) $(A^+)^T = (A^T)^+$, $(A^+)^+ = A$, $(A^+ A)^T = (A^+ A)$, $(AA^+)^T = AA^+$; 3) $AA^+A = A$, $A^+AA^+ = A^+$, $(A^+A)^2 = A^+A$, $(AA^+)^2 = AA^+$; 4) *если U — ортогональная матрица*, $\dim U = m \times m$, *то* $(UA)^+ = A^+ U^T$; 5) *для любого числа* $\lambda \neq 0$ $(\lambda A)^+ = \lambda^{-1} A^+$.

В дальнейшем мы будем использовать евклидовы нормы матриц. Если матрица A имеет размер $m \times n$, то ее евклидова норма вычисляется по формуле

$$\| A \| = \left\{ \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right\}^{1/2}.$$

Евклидовы нормы для матриц размера $m \times m$, $n \times n$ будем обозначать соответственно как $\| \cdot \|_m$, $\| \cdot \|_n$.

Евклидова норма матрицы не меняется при умножении матрицы слева или справа на любую ортогональную матрицу.

Теорема 3 (экстремальное свойство псевдообратной матрицы [49]). *Для любой матрицы Z размера $n \times m$ выполнено неравенство*

$$\| AA^+ - E \|_m \leqslant \| AZ - E \|_m,$$

где E — единичная матрица размера $m \times m$. При этом, если $\| AA^+ - E \|_m = \| AZ - E \|_m$, но $Z \neq A^+$, то $\| A^+ \| < \| Z \|$.

Пусть I — единичная матрица размера $n \times n$. Справедлива

Теорема 4 (псевдообратная матрица как предел [7]). *Имеют место соотношения*

$$A^+ = \lim_{\alpha \to +0} (\alpha I + A^T A)^{-1} A^T, \quad A^+ = \lim_{\alpha \to +0} A^T (\alpha E + AA^T)^{-1}.$$

Обозначим символом $\operatorname{Pr}_L x$ ортогональную проекцию некоторого вектора x на подпространство L.

Теорема 5 (об ортогональном проектировании). *Для любых $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$ справедливы равенства*

$$\operatorname{Pr}_{\operatorname{Im} A} y = AA^+ y, \quad \operatorname{Pr}_{\operatorname{Im} A^T} x = A^+ A x.$$

Теорема 6. *Всякое псевдорешение $\tilde{z} \in \mathbf{R}^n$ системы* (1) *представимо в виде*

$$\tilde{z} = A^+ u + (I - A^+ A) w,$$

где w — некоторый элемент из $\mathbf{R}^n$.

Обратимся теперь к важному понятию теории матриц — сингулярному разложению.

Теорема 7. *Любую матрицу A размера $m\times n$ можно представить в виде $A=URV^T$, где U, V — ортогональные матрицы соответственно размера $m\times m$, $n\times n$, а $R=\operatorname{diag}(\rho_1, \ldots, \rho_M)$ — прямоугольная диагональная матрица размера $m\times n$, содержащая на диагонали неотрицательные числа $\rho_1, \ldots, \rho_M$ $(M=\min(m, n))$, которые упорядочены по невозрастанию*: $\rho_1 \geqslant \ldots \geqslant \rho_M \geqslant 0$.

Числа ρ_k $(k=1, \ldots, M)$ называются *сингулярными числами* матрицы A (A^T). Числа ρ_k^2 $(k=1, \ldots, M)$ являются собственными значениями матриц A^TA, AA^T. Столбцы матриц U, V образуют ортонормированный базис из собственных векторов соответственно матриц AA^T, A^TA. Разложение $A=URV^T$ называется *сингулярным разложением* матрицы A.

Обозначим символом $\operatorname{Rg}A$ ранг матрицы A. Тогда $\operatorname{Rg}A=\operatorname{Rg}R$, и поэтому $\rho_k>0$ при $1\leqslant k\leqslant \operatorname{Rg}A$ и $\rho_k=0$ при $k>\operatorname{Rg}A$.

Справедливы следующие формулы:

$$R^+=\operatorname{diag}(\rho_1^{-1}, \ldots, \rho_r^{-1}, 0, \ldots, 0), \quad r\equiv \operatorname{Rg}A; \quad A^+=VR^+U^T.$$

Кроме того,

$$\|A\|^2=\|R\|^2=\sum_{k=1}^{M}\rho_k^2=\sum_{k=1}^{r}\rho_k^2; \quad \|A^+\|^2=\|R^+\|^2=\sum_{k=1}^{r}\rho_k^{-2}.$$

Если A, $\tilde{A}$ — матрицы размера $m\times n$, а ρ_k, $\tilde{\rho}_k$ — их сингулярные числа, то

$$\|A-\tilde{A}\|^2\geqslant \sum_{k=1}^{M}(\rho_k-\tilde{\rho}_k)^2.$$

Подробное изложение теории всевдообращения матриц можно найти в книгах [7, 46, 143], а также в [238, 240]. Теория сингулярного разложения детально дана в [46, 143].

ГЛАВА 1

ВАРИАЦИОННЫЕ МЕТОДЫ РЕШЕНИЯ НЕКОРРЕКТНЫХ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

§ 1. Корректность постановки экстремальных задач

Многие важные прикладные задачи с математической точки зрения могут быть сведены к поиску экстремума некоторого функционала при определенных ограничениях на его аргумент. Таковыми, например, являются задачи оптимального проектирования технических конструкций, задачи оптимального управления процессами, задачи оптимального планирования и др. Эти задачи можно сформулировать в следующей общей форме.

Пусть (Z, τ) — топологическое пространство и на некотором его подмножестве $D(J)$ определен функционал $J(z)$. Пусть, кроме того, задано непустое множество $D(D \subset D(J))$ и на этом множестве функционал $J(z)$ ограничен снизу. Требуется минимизировать функционал $J(z)$ на множестве D.

Различают два типа задач минимизации. *Задачи первого типа* — это те, в которых необходимо найти минимальное значение функционала J на D: найти число

$$J^* = \inf\{J(z):\ z \in D\}. \tag{1}$$

При этом не имеет значения, на каких элементах из множества D достигается точная нижняя грань и достигается ли она вообще. Задачи первого типа называют также задачами *минимизации по функционалу*.

Задачи второго типа связаны с нахождением элементов множества D, реализующих точную нижнюю грань (1): найти элементы $z^* \in D$, для которых

$$J(z^*) = \inf\{J(z):\ z \in D\} \equiv J^*. \tag{2}$$

Такие задачи называются задачами *минимизации по аргументу*.

Множество решений задачи (2) записывают в форме

$$Z^* \equiv \operatorname{Arg\,inf}\{J(z):\ z \in D\}.$$

Как правило, при практическом решении задач типа (1), (2) точное их решение не удается найти и для нахождения приближенного решения приходится использовать численные методы. Эти численные методы обычно связаны с построением минимизирующих последовательностей для функционала J на множестве D, т. е. таких последовательностей $\{z_n\} \subset D$, что $J(z_n) \to J^*$ при $n \to \infty$. Методы построения минимизирующих последовательностей разработаны в настоящее время для широких классов функционалов. Имеется обширная литература по этому вопросу, из которой упомянем, например, обзорные и итоговые работы [21, 30, 31, 91, 145, 170—173, 212, 216]. Использование этих методов, однако, предполагает, как правило, что задачи (1), (2) обладают важным свойством — корректностью постановки.

Напомним, что математическая задача называется *корректно поставленной* или *корректной*, если выполнены три условия [229]:

1) задача разрешима на заданном классе предполагаемых решений;

2) решение ее единственно на этом классе;

3) решение устойчиво на этом классе по отношению к допустимым возмущениям данных задачи.

При нарушении хотя бы одного из этих требований задача называется *некорректно поставленной.*

Конкретизируем это общее понятие применительно к задачам минимизации (1), (2).

Пусть $\delta \in \mathbf{R}^k$ — конечномерный вектор с неотрицательными координатами и с нормой $\|\delta\|$. Рассмотрим некоторое множество $\mathscr{F}$ функционалов $J_\delta(z)$, зависящих от векторного параметра δ: $0 \leqslant \|\delta\| < \Delta_0 = \text{const}$ и удовлетворяющих следующим условиям:

а) при каждом δ $(0 \leqslant \|\delta\| < \Delta_0)$ все функционалы $J_\delta(z)$ из семейства $\mathscr{F}$ определены на D;

б) функционал $J(z)$ содержится в $\mathscr{F}$, причем он и только он соответствует значению параметра $\delta = (0, ..., 0) \in \mathbf{R}^k$;

в) для любого $z \in D$ выполнено предельное соотношение $\lim\limits_{\delta \to 0} J_\delta(z) = J(z)$ независимо от выбора функционала $J_\delta(z)$ из $\mathscr{F}$ для каждого δ $(0 \leqslant \|\delta\| < \Delta_0)$.

Будем называть такое множество $\mathscr{F}$ *семейством допустимых приближенных функционалов* для задач (1), (2). Подмножество $\mathscr{F}_\Delta$ семейства $\mathscr{F}$, состоящее из функционалов $J_\delta(z)$, для которых дополнительно требуется выполнение неравенства $\|\delta\| \leqslant \Delta$ (Δ — фиксированное число, $0 \leqslant \Delta < \Delta_0$), назовем *семейством допустимых приближенных функционалов, отвечающих уровню погрешности* Δ.

Очевидно, можно считать, что $\mathscr{F} = \bigcup\limits_{0 \leqslant \Delta < \Delta_0} \mathscr{F}_\Delta$. Различные способы задания семейств $\mathscr{F}_\Delta$ будут указаны ниже.

Предположим, что для задач (1), (2) определены семейства $\mathscr{F}$, $\mathscr{F}_\Delta$ $(0 \leqslant \Delta < \Delta_0)$.

Определение 1. Задача минимизации по функционалу (1) называется *корректно поставленной на множестве* $\mathscr{F}$ допустимых приближенных функционалов, если число J^* обладает следующим свойством устойчивости: для любой последовательности $\{\delta_n\}$, сходящейся к нулю при $n \to \infty$, и для любой соответствующей последовательности $\{J_{\delta_n}(z)\}\subset\mathscr{F}$ допустимых приближенных функционалов числовая последовательность $\{J^*_{\delta_n}\}$:

$$J^*_{\delta_n}=\inf\{J_{\delta_n}(z):\ z\in D\}$$

сходится при $n\to\infty$ к J^*.

Проиллюстрируем это определение.

Пример 1. Пусть $Z=\mathbf{R}$, $D=\{z\in\mathbf{R}:\ z\geqslant 0\}$ и $J(z)\equiv 0$. Топология τ определяется евклидовой метрикой в пространстве $\mathbf{R}$: $\rho(z_1, z_2)=|z_1-z_2|$. Ясно, что $J^*=0$. Устойчивость величины J^* определяется классом допустимых приближенных функционалов для $J(z)$. Если, например, считать $\delta=\Delta$ и задать множество $\mathscr{F}_\Delta$ в виде $\mathscr{F}_\Delta=\{J_\delta(z):\ |J(z)-J_\delta(z)|\leqslant\delta\}$ $(0\leqslant\delta<\Delta_0)$, то из неравенства $J(z)-\delta\leqslant J_\delta(z)\leqslant J(z)+\delta$ следует, что $J^*-\delta\leqslant J^*_\delta\leqslant J^*+\delta$, и поэтому $J^*_\delta\to J^*$ при $\delta\to 0$. В этом случае задача (1) оказывается корректно поставленной на классе $\mathscr{F}=\bigcup\limits_{0\leqslant\Delta<\Delta_0}\mathscr{F}_\Delta$ допустимых приближенных функционалов. Если же принять $\mathscr{F}_\Delta=\{J_\delta(z):\ |J(z)-J_\delta(z)|\leqslant\delta|z|\}$, то, выбрав из $\mathscr{F}_\Delta$ специальное подмножество функционалов вида $J_\delta(z)=-\delta z\exp\{-\delta z\}$, получим $J^*_\delta=-1/e$, и поэтому $J^*_\delta\not\to J^*=0$ при $\delta\to 0$. Таким образом, свойство устойчивости на таком классе $\mathscr{F}$ нарушается, и задача (1) оказывается некорректно поставленной.

Пример 2. Пусть $\bar{A}$, A_h — матрицы размера $m\times n$ такие, что $\|\bar{A}-A_h\|\leqslant h$, $\bar{u}$, $u_\sigma\in\mathbf{R}^m$ — вектор-столбцы, удовлетворяющие неравенству $\|\bar{u}-u_\sigma\|\leqslant\sigma$. Все нормы евклидовы. Рассмотрим задачу нахождения так называемой *меры несовместности* $\bar{\mu}$ системы линейных алгебраических уравнений с расширенной матрицей $(\bar{A}, \bar{u})$. Для этого введем *функционал невязки*

$$J(z)=\|\bar{A}z-\bar{u}\|,\quad z\in Z=\mathbf{R}^n,$$

и будем искать

$$\bar{\mu}=\inf\{\|\bar{A}z-\bar{u}\|:\ z\in\mathbf{R}^n\}=\inf\{J(z):\ z\in Z\}. \tag{3}$$

В качестве допустимых приближенных функционалов для невязки рассмотрим множество $\mathscr{F}=\{J_\delta(z)=\|A_hz-u_\sigma\|\}$ $(\delta\equiv(\sigma, h))$. Задача (3) будет корректно поставленной, если система $(\bar{A}, \bar{u})$ разрешима, т. е. существует такой вектор $\bar{z}\in\mathbf{R}^n$, что $\bar{A}\bar{z}=\bar{u}$. Действительно, в этом случае справедлива оценка

$$\mu_{\sigma h}\equiv\inf\{\|A_hz-u_\sigma\|:\ z\in\mathbf{R}^n\}\leqslant\|A_h\bar{z}-u_\sigma\|\leqslant$$
$$\leqslant\|A_h\bar{z}-\bar{A}\bar{z}\|+\|\bar{A}\bar{z}-\bar{u}\|+\|\bar{u}-u_\sigma\|\leqslant h\,\|\bar{z}\|+\sigma+\bar{\mu}=h\,\|\bar{z}\|+\sigma,$$

и поэтому $\mu_{\sigma h}\to\bar{\mu}=0$ при $h, \sigma\to 0$. Для несовместных систем $(\bar{A}, \bar{u})$ сходимость $\mu_{\sigma h}\to\bar{\mu}$, вообще говоря, отсутствует и задача определения меры несовместимости является некорректно поставленной.

Это наглядно иллюстрируется простейшим частным случаем:

$$\bar{A}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix},\quad A_h=\begin{bmatrix}1 & 0\\ 0 & h\end{bmatrix},\quad \bar{u}=u_\sigma=\begin{bmatrix}1\\ 1\end{bmatrix},\quad z=\begin{bmatrix}x\\ y\end{bmatrix},$$

в котором

$$\mu=\inf\{(x-1)^2+1:\ x\in\mathbf{R}\}=1,\quad \mu_{\sigma h}=\inf\{(x-1)^2+(hy-1)^2:\ (x,y)\in\mathbf{R}^2\}=0.$$

Необходимо подчеркнуть, что при решении возникающих на практике систем линейных алгебраических уравнений с данными, отягощенными погрешностями, информация о совместности «точной системы» берется, как правило, из априорных соображений. В условиях отсутствия такой априорной информации задача (3) будет в общем случае некорректно поставленной.

Корректные задачи (1) минимизации по функционалу с вычислительной точки зрения обладают важной особенностью: для их решения при заданном точном $J(z)$ достаточно простроить любую минимизирующую последовательность $\{z_n\}$ для задачи (1) и в качестве приближения для J^* взять $J(z_n)$ при достаточно большом n. В случае, если вместо J известен лишь приближенный функционал $J_\delta\in\mathscr{F}$, можно аналогичным образом построить минимизирующую последовательность $\{z_n\}$ для $J_\delta(z)$ на множестве D. Если при этом для номера $N\equiv n(\delta)$ выполнено неравенство $|J_\delta(z_N)-J_\delta^*|\leqslant\varepsilon(\delta)$, где $\varepsilon(\delta)\to0$ при $\delta\to0$, то $|J_\delta(z_N)-J^*|\leqslant|J_\delta(z_N)-J_\delta^*|+|J_\delta^*-J^*|\leqslant\varepsilon(\delta)+|J_\delta^*-J^*|\to0$, и, следовательно, величину $J_\delta(z_{n(\delta)})$ можно принять в качестве устойчивого приближения к J^*. Для некорректных задач подобный подход неприемлем.

Определение 2. Задача второго типа обладает *свойством устойчивости* на классе $\mathscr{F}$ допустимых приближенных функционалов, если для любой последовательности $\{\delta_n\}$, сходящейся к нулю при $n\to\infty$, и для соответствующей этой $\{\delta_n\}$ произвольной последовательности $\{J_{\delta_n}(z)\}\subset\mathscr{F}$ приближенных функционалов множества

$$Z_{\delta_n}^*\equiv\operatorname{Arg}\inf\{J_{\delta_n}(z):\ z\in D\}$$

не пусты при $n\geqslant n_0=\mathrm{const}$ и сходятся в топологии τ к множеству Z^* решений задачи (2) при $n\to\infty$.

Определение 3. Задача второго типа называется *корректно поставленной*, если ее решение $z^*\in D$ существует, единственно и эта задача обладает свойством устойчивости.

Естественно, что корректность постановки задачи (2) зависит от выбранной топологии τ и от класса $\mathscr{F}$.

Многие задачи второго типа, встречающиеся в приложениях, оказываются некорректно поставленными, так как их решение может не существовать, быть не единственным и не обладать устойчивостью в заданном классе $\mathscr{F}$. Это справедливо, например, для задач оптимального управления [195, 203], задач синтеза антенн [207], многослойных оптических покрытий [210], задач линейного программирования [208] и ряда других задач.

Пример 3. Зададим (Z,τ) и D, как в примере 1, и положим $J(z)=\{z:\ 0\leqslant z\leqslant1;\ 1/z:\ z>1\}$ (рис. 1). Ясно, что единственным решением задачи (2) будет $z^*=0$. Зададим теперь класс допустимых приближенных функционалов, обеспечивающих равномерную аппроксимацию J на D: $\mathscr{F}_\Delta\equiv\{J_\delta:\ |J_\delta-J|\leqslant\delta\}$, где $\Delta=\delta$. Выберем из $\mathscr{F}_\Delta$ конкретные приближенные функционалы: $J_\delta(z)=\{(1-\delta)z:\ 0\leqslant z\leqslant1;\ 1/z:\ 1<z<1/\delta;$

$-\delta$: $z \geqslant 1/\delta\}$. Тогда $Z_\delta^* = \{z \geqslant 1/\delta\}$ и $J_\delta^* = -\delta$, так что $J_\delta^* \to J^* = 0$ при $\delta \to 0$, но Z_δ^* не сходится в топологии, порождаемой евклидовой метрикой, к $z^* = 0$. В данном случае задача (2) оказывается неустойчивой.

Необходимо отметить, что даже для корректной задачи (2) проблема нахождения приближений к z^* не элементарна, так как,

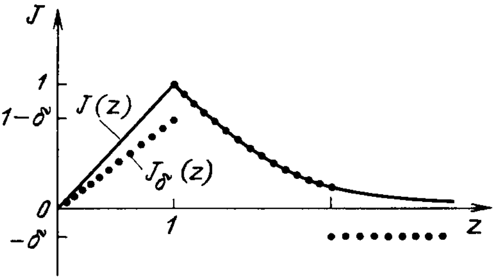

Рис. 1. Иллюстрация к примеру 3

вообще говоря, не всякая минимизирующая последовательность для приближенного функционала J_δ на D будет сходящейся.

Определение 4. Задача второго типа называется *усиленно корректной*, если наряду с условиями определения 3 для нее справедливо условие: для любого $J_\delta \in \mathscr{F}$ $(0 \leqslant \|\delta\| \leqslant \Delta_0)$ всякая минимизирующая последовательность экстремальной задачи

$$\inf\{J_\delta(z):\ z \in D\} \tag{4}$$

τ-сходится к Z_δ^*.

Приведем пример корректной задачи, не являющейся усиленно корректной.

Пример 4. Снова (Z, τ) и D задаются, как в примере 1, а $J(z) = \{z^2:\ 0 \leqslant z \leqslant 1;\quad 1:\ z \geqslant 1\}$. Определим множество допустимых функционалов для данного уровня погрешности $\Delta = \delta$: $\mathscr{F}_\Delta = \{J_\delta(z):\ |J_\delta(z) - J(z)| \leqslant \leqslant \delta(1 + |z|)\}$ и выделим из $\mathscr{F}_\Delta$ функционалы вида $J_\delta(z) = = \{z^2:\ 0 \leqslant z \leqslant 1;\quad 1 - 2\pi^{-1}\operatorname{arctg}[\delta \times \times (z-1)]:\ z \geqslant 1\}$. Для таких J_δ, как легко видеть, $Z_\delta^* = \{0\} = z^*$, но не любая минимизирующая последовательность будет сходящейся (рис. 2).

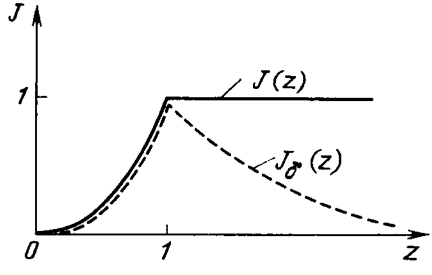

Рис. 2. Иллюстрация к примеру 4

Особенностью усиленно корректных задач типа (2) является возможность приближенного нахождения их решения с помощью минимизирующих последовательностей.

Теорема 1. *Если задача* (2) *усиленно корректная и* $\{z_n(\delta)\}$ — *произвольная минимизирующая последовательность для* (4) *при фиксированном* δ, *то* $z_n(\delta)\xrightarrow{\tau}z^*$ *при* $n\to\infty$, $\delta\to 0$.

Доказательство. По определению 3 имеют место сходимости: $z_n(\delta)\xrightarrow{\tau}Z_\delta^*$ при $n\to\infty$ и фиксированном δ $(0\leqslant\delta\leqslant\delta_0)$; $Z_\delta^*\xrightarrow{\tau}z^*$ при $\delta\to 0$. Из второй сходимости следует, что для любой окрестности $G\subset\tau$ точки z^* найдется число $\delta_1(G)$ такое, что при $0\leqslant\delta\leqslant\delta_1(G)$ справедливо включение $Z_\delta^*\subset G$. Таким образом, при $0\leqslant\delta\leqslant\delta_1$ G есть окрестность множества Z_δ^*. Для этой окрестности, согласно определению сходимости $z_n(\delta)\xrightarrow{\tau}Z_\delta^*$ при $n\to\infty$ и фиксированном δ $(0\leqslant\delta\leqslant\delta_1)$, получим, что существует номер $N(G)$, для которого при любом $n(n>N)$ выполняется включение $z_n(\delta)\in G$. Но тогда по определению τ-сходимости $z_n(\delta)\xrightarrow{\tau}z^*$ при $n\to\infty$, $\delta\to 0$. Теорема доказана.

Из сказанного ясно, что для устойчивого приближенного решения экстремальных некорректных задач требуются специальные методы, называемые обычно *регуляризующими алгоритмами.*

Определение 5. *Регуляризующим алгоритмом для решения задачи* (1) называется отображение I, ставящее в соответствие каждому допустимому набору данных (J_δ, δ) число $I_\delta\equiv I(J_\delta, \delta)$, такое, что $I_\delta\to J^*$ при $\delta\to 0$.

Определение 6. *Регуляризующим алгоритмом для решения задачи* (2) назовем отображение R, задающее для каждых допустимых (J_δ, δ) элемент $z_\delta=R(J_\delta, \delta)\in D$ такой, что $z_\delta\xrightarrow{\tau}Z^*$ при $\delta\to 0$.

В § 2 будет выделен специальный класс некорректных экстремальных задач первого и второго типа, наиболее часто встречающихся на практике.

§ 2. Основные задачи и предположения

Уточним постановку задач минимизации функционалов. Снова будем считать, что (Z, τ) — топологическое пространство, $D(J)$ — его подмножество, являющееся областью определения функционала $J(z)$. Предположим также, что задано множество $D\subset D(J)$, на котором $J(z)$ ограничен снизу. Сформулируем задачи минимизации.

Задача минимизации по функционалу. Найти число

$$J^*=\inf\{J(z):\ z\in D\}. \tag{1}$$

Задача минимизации по аргументу. Найти все элементы $z^*\in D$ такие, что

$$J(z^*)=\inf\{J(z):\ z\in D\}=J^*. \tag{2}$$

При этом будем предполагать, что множество Z^* таких элементов z^* непусто.

Предположим теперь, что точный функционал J неизвестен, а вместо него в нашем распоряжении имеются некоторые «приближенные» функционалы.

Зададим класс допустимых приближенных функционалов. Для этого допустим, что на множестве D задан некоторый вспомогательный функционал $\Omega(z)$, причем он ограничен на D снизу:

$$\Omega^* \equiv \inf\{\Omega(z):\ z\in D\} > -\infty.$$

Пусть $\delta\in\mathbf{R}^k$ — вектор с неотрицательными координатами и с евклидовой нормой. Будем говорить, что функционал $\tilde{J}(z)$, определенный на D и ограниченный там снизу, является *приближенным* с погрешностью $\|\delta\|$ $(0\leqslant\|\delta\|<\Delta_0)$ для функционала $J(z)$, если он удовлетворяет условию аппроксимации

$$|J(z)-\tilde{J}(z)|\leqslant\Psi(\delta, \Omega(z))\quad \forall z\in D. \tag{3}$$

При этом функция Ψ — мера аппроксимации точного функционала $J(z)$ приближенными — считается заданной и обладает рядом свойств, перечисленных в конце параграфа. Вектор δ характеризует по формуле (3) «близость» функционалов J и $\tilde{J}$. Его норма может служить числовой характеристикой такой близости. Таким образом, задание приближенного функционала $\tilde{J}$ предполагает знание величин δ, $\Psi(\delta, \Omega(z))$.

Будем обозначать приближенные функционалы, отвечающие погрешности $\|\delta\|$, как $J_\delta(z)$. Совокупность всех таких функционалов для δ $(0\leqslant\|\delta\|<\Delta_0)$ обозначим как $\mathcal{F}$. При выполнении условий на меру аппроксимации, перечисленных в конце параграфа, это множество $\mathcal{F}$ будет классом допустимых приближенных функционалов (см. § 1).

Итак, здесь и в дальнейшем задание множества $\mathcal{F}$ будет осуществляться путем задания определенной функции $\Psi(\delta, \Omega)$. Приведем простейший пример множества $\mathcal{F}$. Он связан со случаем равномерного приближения функционала J функционалом J_δ на множестве D. В этом случае $\Psi(\delta, \Omega(z))=\Psi(\delta)$, т. е. мера аппроксимации не зависит от элемента $z\in D$, на котором эта аппроксимация производится. Такая аппроксимация возможна лишь для узкого класса данных $\{J, J_\delta, D\}$. Общие требования к мере аппроксимации Ψ приведены в конце параграфа. Конкретные примеры функций $\Psi(\delta, \Omega)$ будут даны в § 11.

Задачи (1), (2) в общем случае являются некорректно поставленными на классе допустимых приближенных функционалов $\mathcal{F}$.

Основная задача первого типа, рассматриваемая ниже, заключается в построении по совокупности приближенных данных (J_δ, δ, Ψ) $(J_\delta\in\mathcal{F})$ задачи (1) числа λ_δ, обладающего тем свойством, что $\lambda_\delta\to J^*$ при $\|\delta\|\to 0$.

Перейдем к задаче (2). Как правило, множество ее решений Z^* содержит более одного элемента. Поэтому возникает проблема выбора из Z^* некоторого элемента, обладающего дополнительными оптимальными свойствами. Математически это выражается в поиске так называемых *Ω-оптимальных решений* задачи (2): найти такие элементы $\bar{z}\in Z^*$, для которых

$$\Omega(\bar{z})=\inf\{\Omega(z):\ z\in Z^*\}\equiv\bar{\Omega}. \tag{4}$$

Будем предполагать, что множество $\bar{Z}=\{\bar{z}\}$ Ω-оптимальных решений задачи (2) непусто.

Задача поиска Ω-оптимальных решений также, вообще говоря, некорректно поставлена. В частности, $\bar{Z}$ может содержать более одного элемента.

Основная задача второго типа, изучаемая в данной книге, связана с нахождением по совокупности (J_δ, δ, Ψ) $(J_\delta \in \mathscr{F})$ приближенных данных задачи (2) элемента $z_\delta \in D$, τ-секвенциально сходящегося при $\|\delta\| \to 0$ к множеству $\bar{Z}$ Ω-оптимальных решений задачи (2).

Рассмотрим несколько примеров задач нахождения Ω-оптимальных решений.

Пример 1. Пусть $\bar{A}$ — матрица размерности $m \times n$, $\bar{u} \in \mathbf{R}^m$ — вектор-столбец, и пусть изучается система линейных алгебраических уравнений

$$\bar{A}z=\bar{u}, \quad z \in \mathbf{R}^n. \tag{5}$$

Эта система в принципе может быть несовместной, но всегда она имеет некоторые обобщенные решения, называемые псевдорешениями. *Псевдорешения* системы (5) определяются как элементы $z^* \in \mathbf{R}^n$, минимизирующие невязку (см. § 0.6)

$$\|\bar{A}z^*-\bar{u}\|=\inf\{\|\bar{A}z-\bar{u}\|:\ z \in \mathbf{R}^n\}. \tag{6}$$

Все нормы считаются для определенности евклидовыми. В частности, если система (5) разрешима, то она эквивалентна экстремальной задаче (6). Таким образом, оставляя в стороне вопросы исследования системы (5) на совместность, можно ставить для нее задачу о поиске псевдорешений в форме (6), т. е. рассматривать задачу типа (2) с $J(z)=\|\bar{A}z-\bar{u}\|$.

Предположим теперь, что вместо точных данных $(\bar{A}, \bar{u})$ задачи (5) заданы приближенные данные (A_h, u_σ), удовлетворяющие требованиям $\|\bar{A}-A_h\| \leqslant h$, $\|\bar{u}-u_\sigma\| \leqslant \sigma$. При этом числа h, $\sigma \geqslant 0$, определяющие уровни погрешности данных, известны. Тогда можно считать, что вместо функционала $J(z)$ задачи (6) задан приближенный функционал $J_\delta(z)=\|A_h z-u_\sigma\|$, где $\delta=(h, \sigma)$, и выполнено условие аппроксимации

$$\begin{aligned}|J(z)-J_\delta(z)| &= |\,\|\bar{A}z-\bar{u}\|-\|A_h z-u_\sigma\|\,| \leqslant \\ &\leqslant \|\bar{A}z-A_h z\|+\|\bar{u}-u_\sigma\| \leqslant h\,\|z\|+\sigma \equiv \Psi(\delta, \|z\|).\end{aligned} \tag{7}$$

Таким образом, задается класс допустимых приближенных функционалов $\mathscr{F}$ для задачи (6), а также мера аппроксимации Ψ. В качестве $\Omega(z)$ здесь принято $\Omega(z)=\|z\|$.

Можно поставить задачу нахождения среди псевдорешений системы (5) такого, которое обладает минимальной нормой (*нормальное псевдорешение*). Это соответствует нахождению Ω-оптимального решения задачи (6) для $\Omega(z)=\|z\|$. Известно, что такое нормальное псевдорешение $\bar{z}$ существует и единственно для любой системы (5), однако оно неустойчиво по отношению к возмущениям данных системы — погрешностям в задании матрицы (см., например, [4, 44, 143, 154, 167], а также гл. 5). Тем самым основная задача второго типа для рассматриваемого примера заключается в построении по приближенным данным $(A_h, u_\sigma, h, \sigma, \Psi)$ системы (5) элемента $z_\delta \in \mathbf{R}^n$, сходящегося при $\delta=(h, \sigma)\to 0$ к нормальному псевдорешению $\bar{z}$.

Пример 2. *Вариационный подход к решению нелинейных операторных уравнений.* Предположим, что Z — нормированное пространство с топологией, порождаемой нормой $\|\cdot\|_Z$. Будем считать, что на множестве $D(D\subset Z)$ определен оператор A (в общем случае нелинейный) со значениями в метрическом пространстве U, которое имеет метрику ρ. Рассмотрим операторное уравнение

$$Az=u, \quad z\in D, \quad u\in U. \tag{8}$$

Это уравнение может и не иметь решений на множестве D. Однако во многих случаях оно имеет так называемые *квазирешения*, представляющие собой решения экстремальной задачи: найти элемент $z^*\in D$ такой, что

$$\rho(Az^*, u)=\inf\{\rho(Az, u)\colon\ z\in D\}. \tag{9}$$

В частном случае разрешимого на D уравнения (8) квазирешения совпадают с обычными решениями (8). Таким образом, задача (9) поиска квазирешений на множестве D оказывается частным случаем задачи (2) с $J(z)=\rho(Az, u)$. Вопросы существования квазирешений на множествах D различной структуры подробно исследованы в работах [77, 78, 84—88, 187, 188]. Если предположить, что множество Z^* квазирешений уравнения (8) на множестве D непусто, то можно поставить задачу о нахождении *нормального квазирешения*, т. е. такого $\bar{z}\in Z^*$, для которого

$$\|\bar{z}\|_Z=\inf\{\|z^*\|_Z\colon\ z^*\in Z^*\}.$$

Задача о нахождении $\bar{z}$ есть частный случай задачи (4) определения Ω-оптимальных решений задачи (9) с $\Omega(z)=\|z\|_Z$. Такая задача часто оказывается некорректно поставленной. Соответствующие примеры даны, например, в [88, 160, 202]. Более подробно вариационный подход к решению нелинейных операторных уравнений рассмотрен в гл. 2.

Пример 3. *Простейшая задача оптимального проектирования конструкций.* Пусть некоторая техническая система характеризуется функцией $z(s)$. Качество такой системы будем характеризовать некоторым функционалом $J[z(s)]$. На функцию $z(s)$ обычно заданы ограничения двух типов: качественные, когда определяется функциональный класс Z, которому принадлежит $z(s)$, и количественные, часто выражаемые в виде некоторых неравенств. Определяемое таким образом множество допустимых функций z обозначим как D. Тогда задача поиска системы, оптимальной в смысле минимума функционала $J(z)$, имеет вид (2). Может оказаться, что таких оптимальных систем несколько. Тогда, задавая дополнительный функционал $\Omega(z)$, характеризующий некоторое другое качество системы, можно, решая задачу типа (4), отобрать из оптимальных по функционалу J такие функции z, которые будут оптимальны еще и по Ω. Конкретные примеры задач такого вида содержатся, например, в работах [95, 107, 127, 137, 174, 175, 202, 207, 210] (см. также гл. 6).

К задачам поиска Ω-оптимальных решений относятся также многие задачи оптимального управления [195], оптимального планирования [196, 198] и ряд других.

В заключение параграфа сформулируем основные предположения относительно функционалов J, J_δ, Ω, а также относительно меры аппроксимации Ψ, которые будут использоваться в данной главе.

1. Функционал $J(z)$ τ-секвенциально полунепрерывен снизу на D: для любой последовательности $\{z_n\}\subset D$, τ-сходящейся при $n\to\infty$ к $z_0\in D$, выполнено неравенство

$$\varliminf_{n\to\infty} J(z_n)\geqslant J(z_0).$$

Аналогичным свойством должны обладать и функционалы J_δ, Ω.

2. Если множество $\Omega_K\equiv\{z\in D\colon \Omega(z)\leqslant K\}$, где $K=\mathrm{const}$, не пусто, то оно τ-секвенциально компактно.

3. Функция $\Psi(\delta, \Omega)$ определена для всех $\delta\in\mathbf{R}^k_+$, т. е. для $\delta=(\delta_1, ..., \delta_k)\in\mathbf{R}^k$, где $\delta_i\geqslant 0$ $(i=1, ..., k)$, и для всех $\Omega\geqslant\Omega^*$. Кроме того, функция Ψ обладает следующими свойствами:

а) Ψ неотрицательна и непрерывна в области определения;

б) $\Psi(\delta, \Omega)$ монотонно не убывает по Ω при каждом допустимом фиксированном δ;

в) $\Psi(\theta, \Omega)=0$ при каждом $\Omega\geqslant\Omega^*$; здесь $\theta=(0, ..., 0)\in\mathbf{R}^k$.

Какие-либо дополнительные свойства величин J, J_δ, Ω, Ψ, отличные от свойств 1—3 и используемые при доказательстве тех или иных утверждений, будут каждый раз оговариваться.

В дальнейшем функционал Ω, обладающий свойствами 1, 2, будем называть *стабилизирующим в топологии* τ (или *стабилизатором*).

§ 3. Устойчивый метод решения задач первого типа

Рассмотрим основную задачу первого типа: по совокупности приближенных данных (J_δ, δ, Ψ) $(J_\delta\in\mathscr{F})$ задачи (2.1) найти такие числа λ_δ, что $\lambda_\delta\to J^*$ при $\|\delta\|\to 0$. Для решения введем новый функционал

$$L_\delta(z)\equiv J_\delta(z)+\Psi(\delta, \Omega(z)),\quad z\in D. \tag{1}$$

Лемма 1. *Функционал* (1) *τ-секвенциально полунепрерывен снизу на D.*

Доказательство. Действительно, пусть $\{z_n\}\subset D$ — произвольная τ-сходящаяся к $z_0\in D$ последовательность. Тогда из основного предположения 2.1 следует, что

$$\varliminf_{n\to\infty} J_\delta(z_n)\geqslant J_\delta(z_0),\quad \varliminf_{n\to\infty}\Omega(z_n)\geqslant\Omega(z_0).$$

В силу монотонного неубывания функции Ψ по второму аргументу и в силу ее непрерывности

$$\varliminf_{n\to\infty}\Psi(\delta, \Omega(z_n))\geqslant\Psi(\delta, \Omega(z_0)).$$

Полученные соотношения вместе с (1) дают

$$\varliminf_{n\to\infty} L_\delta(z_n)=\varliminf_{n\to\infty}[J_\delta(z_n)+\Psi(\delta,\Omega(z_n))]\geqslant$$

$$\geqslant \varliminf_{n\to\infty} J_\delta(z_n)+\varliminf_{n\to\infty}\Psi(\delta,\Omega(z_n))\geqslant J_\delta(z_0)+\Psi(\delta,\Omega(z_0))=L_\delta(z_0),$$

что и требовалось доказать.

Будем рассматривать величины $J_\delta^*=\inf\{J_\delta(z)\colon z\in D\}$, и

$$\lambda_\delta=\inf\{L_\delta(z)\colon z\in D\} \tag{2}$$

(см. [124, 125]).

Как показано на примерах в § 1 числа J_δ^*, вообще говоря, не могут использоваться в качестве устойчивого приближения к J^*. Оказывается, что числа λ_δ обладают необходимым свойством устойчивости.

Теорема 1. 1) *Справедливы неравенства* $J^*\leqslant\lambda_\delta$, $J_\delta^*\leqslant\lambda_\delta$.

2) *Если* $\|\delta\|\to 0$, *то* $\lambda_\delta\to J^*$.

Доказательство. Используем неравенство (2.3), определяющее класс $\mathscr{F}$ допустимых приближенных функционалов J_δ. Из (2.3) получим

$$J^*\leqslant J(z)\leqslant J_\delta(z)+\Psi(\delta,\Omega(z))=L_\delta(z)\quad \forall z\in D, \tag{3}$$

и поэтому $J^*\leqslant\inf\{L_\delta(z)\colon z\in D\}=\lambda_\delta$. Далее, учитывая неотрицательность функции Ψ, получаем

$$J_\delta^*\leqslant J_\delta(z)\leqslant J_\delta(z)+\Psi(\delta,\Omega(z))=L_\delta(z)\quad \forall z\in D,$$

откуда $J_\delta^*\leqslant\inf\{L_\delta(z)\colon z\in D\}=\lambda_\delta$.

Для доказательства второй части теоремы напишем вытекающее из (3), (2.3) неравенство

$$J^*\leqslant\lambda_\delta\leqslant J_\delta(z)+\Psi(\delta,\Omega(z))\leqslant J(z)+2\Psi(\delta,\Omega(z))\quad \forall z\in D. \tag{4}$$

Затем выберем произвольную минимизирующую последовательность $\{z_n\}\subset D$ для задачи (2.1): $J(z_n)\to J^*$ при $n\to\infty$. Тогда для любого $\varepsilon>0$ найдется номер $N=N(\varepsilon)$ такой, что $0\leqslant J(z_N)-J^*<\varepsilon$. Это неравенство вместе с (4) дает

$$J^*\leqslant\lambda_\delta\leqslant J(z_N)+2\Psi(\delta,\Omega(z_N))<J^*+\varepsilon+2\Psi(\delta,\Omega(z_N)).$$

Переходя в последнем неравенстве к пределу при $\delta\to 0$ и учитывая, что, согласно свойствам функции Ψ, $\Psi(\delta,\Omega(z_N))\to 0$ при $\delta\to 0$, получаем

$$J^*\leqslant\varliminf_{\delta\to 0}\lambda_\delta\leqslant\varlimsup_{\delta\to 0}\lambda_\delta\leqslant J^*+\varepsilon\quad \forall\varepsilon>0.$$

В силу произвольности ε отсюда следует, что $\lambda_\delta\to J^*$ при $\|\delta\|\to 0$.

Согласно теореме 1, величины λ_δ могут служить устойчивыми приближениями для J^* в классе $\mathscr{F}$ приближенных функционалов, а всякая минимизирующая последовательность $\{z_m\}\subset D$ для задачи (2) при достаточно большом m и достаточно малом $\|\delta\|$ дает приближение $L_\delta(z_m)$ для J^* (см. § 1).

Пример 1. Рассмотрим задачу об устойчивом вычислении меры несовместности $\bar{\mu}=\inf\{\|\bar{A}z-\bar{u}\|:\ z\in\mathbf{R}^n\}$ системы линейных алгебраических уравнений $\bar{A}z=\bar{u}$ (см. пример 1.2). Здесь $\bar{A}$ — матрица размера $m\times n$, $\bar{u}\in\mathbf{R}^m$ и все нормы евклидовы. Приближенными данными служат матрица A_h того же размера и элемент $u_\sigma\in\mathbf{R}^m$, для которых справедливы оценки $\|\bar{A}-A_h\|\leqslant h$, $\|\bar{u}-u_\sigma\|\leqslant\sigma$. Число h, σ предполагаются известными. Вводя $\Omega(z)=\|z\|$, можно считать, что $\Psi(\delta,\Omega(z))=$ $=\sigma+h\|z\|$ $(\delta\equiv(h,\sigma))$. Тогда по теореме 1 в качестве устойчивого приближения к $\bar{\mu}$ можно взять число $\lambda_\delta=\inf\{\|A_hz-u_\sigma\|+h\|z\|+\sigma:\ z\in$ $\in\mathbf{R}^n\}$ — *обобщенную меру несовместности* [39, 103, 105].

Задача (2) обладает также и другими свойствами, облегчающими ее численное решение.

Теорема 2. *Пусть в дополнение к свойствам функции* Ψ (*см.* § 2) *она такова, что* $\Psi(\delta,\Omega)\to+\infty$ *при* $\Omega\to+\infty$ *для любого фиксированного* $\delta\in\mathbf{R}^k_+$ $(\delta\neq\theta)$. *Тогда существует такой элемент* $z_0\in D$, *что*

$$L_\delta(z_0)=\inf\{L_\delta(z):\ z\in D\}=\lambda_\delta \tag{5}$$

и всякая минимизирующая последовательность $\{z_m\}$ *для задачи* (2) τ-*сходится к множеству* $\{z_0\}$ *решений задачи* (5).

Доказательство. Если $\{z_m\}$ — произвольная минимизирующая последовательность для задачи (2), то при $m\geqslant m_0(\varepsilon)$ из теоремы 1 получим

$$J^*_\delta+\Psi(\delta,\Omega(z_m))\leqslant J_\delta(z_m)+\Psi(\delta,\Omega(z_m))\leqslant\lambda_\delta+\varepsilon,$$

и поэтому $\Psi(\delta,\Omega(z_m))\leqslant\lambda_\delta-J^*_\delta+\varepsilon\equiv K_0>0$. Из этого неравенства и условий теоремы следует ограниченность величин $\Omega(z_m)$: $\Omega(z_m)\leqslant\tilde{\Omega}=\text{const}$, и, значит, согласно основному предположению 2 из § 2, последовательность $\{z_m\}$ оказывается τ-секвенциально компактной.

Выделим из $\{z_m\}$ τ-сходящуюся последовательность $\{z_{m_k}\}$: $z_{m_k}\xrightarrow{\tau}z_0\in D$ при $k\to\infty$. Тогда, учитывая, что $\{z_{m_k}\}$ — минимизирующая для задачи (2), а по лемме 1 функционал $L_\delta(z)$ τ-секвенциально полунепрерывен снизу, получаем

$$\lambda_\delta=\lim_{k\to\infty}L_\delta(z_{m_k})\geqslant L_\delta(z_0)\geqslant\lambda_\delta,$$

т. е. $L_\delta(z_0)=\lambda_\delta$ и $z_0\in D$ — решение задачи (5).

Докажем теперь, что вся минимизирующая последовательность $\{z_m\}$ τ-сходится к множеству $\{z_0\}$ решений задачи (5). Пусть это не так, т. е. для некоторой окрестности $G\subset\tau$ множества $\{z_0\}$ и для некоторой последовательности номеров $\{m_p\}_{p=1}^\infty$ оказывается, что $z_{m_p}\notin G$. Вместе с тем выше показано, что множества $\{z_m\}$ и, следовательно, $\{z_{m_p}\}$ — τ-секвенциально компактные, и, значит, из $\{z_{m_p}\}$ можно извлечь подпоследовательность $\{z_l\}$ $(l=m_{p_k},\ k=1,2,\ldots)$, которая аналогичным образом будет τ-сходиться к элементу $z^*\in\{z_0\}$. Но тогда по определению τ-сходимости $z_l\in G$ для всякого номера $l\geqslant N(G)=\text{const}$. Полученное противоречие доказывает сходимость $z_m\xrightarrow{\tau}\{z_0\}$ при $m\to\infty$.

Допустим теперь, что функционал $L_\delta(z)$ сам известен не точно: вместо него задан функционал $L_{\delta\sigma}(z)$ $(z \in D)$, удовлетворяющий условию аппроксимации:

$$|L_\delta(z) - L_{\delta\sigma}(z)| \leqslant \sigma \quad \forall z \in D. \tag{6}$$

Теорема 3. *Пусть* $\lambda_{\delta\sigma} \equiv \inf\{L_{\delta\sigma}(z)\colon z \in D\}$. *Тогда* $\lambda_{\delta\sigma} \to \lambda_\delta$ *при* $\sigma \to 0$.

Доказательство. Запишем два следствия неравенства (6):

$$\lambda_\delta \leqslant L_\delta(z_n) \leqslant L_{\delta\sigma}(z_n) + \varepsilon, \quad \lambda_{\delta\sigma} \leqslant L_{\delta\sigma}(z_m) \leqslant L_\delta(z_m) + \sigma.$$

Здесь $\{z_n\} \subset D$ — минимизирующая последовательность для функционала $L_{\delta\sigma}(z)$ на D: $L_{\delta\sigma}(z_n) \to \lambda_{\delta\sigma}$ при $n \to \infty$, а $\{z_m\} \subset D$ — минимизирующая последовательность для задачи (2): $L_\delta(z_m) \to \lambda_\delta$ при $m \to \infty$. После предельного перехода в этих неравенствах получим $\lambda_\delta \leqslant \lambda_{\delta\sigma} + \sigma$, $\lambda_{\delta\sigma} \leqslant \lambda_\delta + \sigma$, откуда и следует сходимость $\lambda_{\delta\sigma} \to \lambda_\delta$ при $\sigma \to 0$.

Из теорем 2, 3 ясно, что корректность задачи (2) (как задачи минимизации по аргументу) может нарушаться лишь из-за наличия в множестве ее решений $\{z_0\}$ более одного элемента. Если же $\{z_0\}$, состоит лишь из одного элемента, то задача (2) по теоремам 2, 3 оказывается усиленно корректной.

Вопросы конкретной численной реализации предложенной в этом параграфе процедуры приближенного нахождения точной нижней грани J^* рассматриваются в § 14.

Замечание. Предлагаемый подход к приближенному нахождению J^* предложен в работах [124, 125], а для случая вариационного решения операторных уравнений — в [39, 103, 105]. Затем он рассматривался также в работе [115] применительно к случаю линейных операторных уравнений. Другой подход изучался в работе [158].

§ 4. Общая схема решения некорректных экстремальных задач минимизации по аргументу

Рассмотрим основную задачу второго типа: по совокупности (J_δ, δ, Ψ) $(J_\delta \in \mathscr{F})$ приближенных данных задачи (2.2) найти элемент $z_\delta \in D$, τ-секвенциально сходящийся при $\|\delta\| \to 0$ к множеству $\bar{Z}$ Ω-оптимальных решений задачи (2.2).

Предположения на функционалы J, J_δ, Ω, введенные в § 2, гарантируют разрешимость экстремальной задачи (2.4) о поиске Ω-оптимальных решений задачи (2.2).

Теорема 1. *Пусть задача (2.2) имеет непустое множество решений Z^*. Тогда множество $\bar{Z}$ Ω-оптимальных решений также непусто.*

Доказательство. Пусть $\{z_n\} \subset Z^*$ — произвольная минимизирующая последовательность для экстремальной задачи (2.4): $\Omega(z_n) \to \bar{\Omega}$ при $n \to \infty$. Тогда для всех $n \geqslant n_0$ выполнено неравенство $\Omega(z_n) \leqslant \bar{\Omega} + \varepsilon \equiv C_0 = \mathrm{const}$ $(\varepsilon > 0)$, и поэтому $\{z_n\} \subset \Omega_{C_0}$ при $n \geqslant n_0$. Тогда в силу предположения 2 из § 2 из $\{z_n\}$ можно извлечь подпоследовательность $\{z_{n_k}\}$ такую, что $z_{n_k} \xrightarrow{\tau} z_0 \in \Omega_{C_0} \subset D$. Вследствие этого из предположения 1 из § 2 получаем

$$\varliminf_{k\to\infty} J(z_{n_k}) = \varliminf_{k\to\infty} J^* = J^* \geqslant J(z_0),$$

$$\varliminf_{k\to\infty} \Omega(z_{n_k}) = \lim_{k\to\infty} \Omega(z_{n_k}) = \bar{\Omega} \geqslant \Omega(z_0),$$

где учтено, что $z_{n_k} \in Z^*$, т. е. $J(z_{n_k}) = J^*$, и что подпоследовательность (z_{n_k}) также является минимизирующей для задачи (2.4). Первое из полученных соотношений в силу (2.2) дает $J(z_0) = J^*$, т. е. $z_0 \in Z^*$. Из второго соотношения и из (2.4) ясно, что $\Omega(z_0) = \bar{\Omega}$, т. е. $z_0 \in \bar{Z}$. Теорема доказана.

Следствие 1. *Множество $\bar{Z}$ τ-секвенциально компактно.*

Действительно, взяв любую последовательность $\{z_n\}$ элементов множества $\bar{Z} \subset Z^*$, можно, как и при доказательстве теоремы 1, убедиться, что существует подпоследовательность, τ-сходящаяся к некоторому элементу $z_0 \in \bar{Z}$.

Тем самым некорректность постановки задачи поиска Ω-оптимальных решений может быть связана лишь с неединственностью такого решения, а также с неустойчивостью задачи на классе $\mathscr{F}$.

В данной главе принята общая схема решения основной задачи второго типа, основанная на методе регуляризации некорректных экстремальных задач [124, 130, 199, 202]. Пусть $f(x)$ — некоторая заданная вспомогательная функция, определенная на $[x_0, +\infty)$, содержащем все возможные значения функционалов J_δ, J. Введем так называемый *модифицированный сглаживающий функционал* для задачи (2.2):

$$M_\delta^\alpha[z] = \alpha\Omega(z) + I_\delta(z), \quad \alpha > 0, \quad z \in D, \tag{1}$$

где $I_\delta(z) \equiv f[J_\delta(z)]$. Этот функционал будет использоваться при фиксированном δ для различных значений параметра регуляризации $\alpha > 0$. Будем рассматривать при каждом $\alpha > 0$ задачу на условный экстремум для функционала (1): найти элемент $z_\delta^\alpha \in D$ такой, что

$$M_\delta^\alpha[z_\delta^\alpha] = \inf\{M_\delta^\alpha[z]: z \in D\}. \tag{2}$$

Лемма 1. *Пусть $f(x)$ непрерывна и монотонно возрастает на $[x_0, +\infty)$, а для J_δ, Ω выполнено предположение* 1 *из* § 2. *Тогда функционалы $M_\delta^\alpha[z]$, $I_\delta(z)$ τ-секвенциально полунепрерывны снизу на D.*

Доказательство. Пусть $\{z_n\} \subset D$ — произвольная последовательность, τ-сходящаяся к элементу $z_0 \in D$ при $n \to \infty$. Тогда из предположения 1 из § 2 следует, что

$$\varliminf_{n\to\infty} \Omega(z_n) \geqslant \Omega(z_0), \quad \varliminf_{n\to\infty} J_\delta(z_n) \geqslant J_\delta(z_0),$$

откуда с учетом непрерывности и монотонности функции f получим

$$\varliminf_{n\to\infty} f[J_\delta(z_n)] \geqslant f[J_\delta(z_0)],$$

$$\varliminf_{n\to\infty} M_\delta^\alpha[z_n] \geqslant \varliminf_{n\to\infty} \alpha\Omega(z_n) + \varliminf_{n\to\infty} f[J_\delta(z_n)] \geqslant \alpha\Omega(z_0) + f[J_\delta(z_0)] = M_\delta^\alpha[z_0].$$

Это, согласно определению, и означает τ-секвенциальную полунепрерывность снизу на D функционала (1) и функционала $I_\delta(z)$.

Теорема 2. *Пусть $f(x)$ непрерывна и монотонно возрастает на $[x_0, +\infty)$, а для J_δ, Ω выполнены предположения* 1, 2 *из* § 2. *Тогда задача* (2) *разрешима, а множество ее решений Z_δ^α τ-секвенциально компактно.*

Доказательство. Выберем $\tilde{z}$ — некоторый фиксированный элемент из D. Тогда можно ввести множество $\tilde{Z}=\{z\in D: M_\delta^\alpha[z]\leqslant M_\delta^\alpha[\tilde{z}]\}\neq\varnothing$. Для элементов $z\in\tilde{Z}$ выполняется неравенство

$$\Omega(z)\leqslant\{M_\delta^\alpha[\tilde{z}]-f[J_\delta(z)]\}/\alpha\leqslant\{M_\delta^\alpha[\tilde{z}]-f[J_\delta^*]\}/\alpha\equiv P_0,$$

где константа P_0 не зависит от z. Таким образом, $\tilde{Z}\subset\Omega_{P_0}$. Покажем, что $\tilde{Z}$ — τ-секвенциально компактное множество. Действительно, пусть $\{z_n\}$ — произвольная последовательность из $\tilde{Z}$. Тогда $\{z_n\}\subset\Omega_{P_0}$ и в силу τ-секвенциальной компактности множества Ω_{P_0} (см. предположение 2 из § 2) найдется подпоследовательность $\{z_{n_k}\}$ такая, что $z_{n_k}\xrightarrow{\tau}z^*\in\Omega_{P_0}\subset D$. По лемме 1 и с учетом включения $z_{n_k}\in\tilde{Z}$ получим

$$M_\delta^\alpha[\tilde{z}]\geqslant\varliminf_{k\to\infty}M_\delta^\alpha[z_{n_k}]\geqslant M_\delta^\alpha[z^*].$$

Тем самым $z^*\in\tilde{Z}$ и секвенциальная компактность множества $\tilde{Z}$ установлена.

Итак, функционал (1), являясь по лемме 1 τ-секвенциально полунепрерывным снизу на D, а потому и на $\tilde{Z}$, определен на τ-секвенциально компактном множестве $\tilde{Z}$. Поэтому, согласно обобщенной теореме Вейерштрасса (см. § 1 вводной главы), экстремальная задача

$$\operatorname{Arg}\inf\{M_\delta^\alpha[z]: z\in\tilde{Z}\}$$

разрешима, т. е. существует такой элемент $z_\delta^\alpha\in\tilde{Z}$, для которого

$$M_\delta^\alpha[z_\delta^\alpha]=\inf\{M_\delta^\alpha[z]: z\in\tilde{Z}\}=\inf\{M_\delta^\alpha[z]: z\in D\}.$$

Это и означает разрешимость задачи (2). Секвенциальная компактность множества Z_δ^α решений задачи (2) следует из теоремы 2, приведенной в § 1 вводной главы.

Процедура решения основной задачи второго типа использует решения $\{z_\delta^\alpha\}=Z_\delta^\alpha$ задачи (2) и состоит из двух взаимосвязанных частей:

1) выбора параметра регуляризации $\alpha(\delta)$ по приближенным данным $\{I_\delta, \Psi, \delta\}$ задачи (2.2) в соответствии с некоторым принципом,

2) фиксации множества $Z_\delta^{\alpha(\delta)}$ экстремалей функционала (1) для найденного $\alpha=\alpha(\delta)$ и выбора из него экстремали $z_\delta^{\alpha(\delta)}$ по некоторому *правилу отбора*.

Полученный таким образом элемент $z_\delta^{\alpha(\delta)}\in D$ принимается в качестве приближенного решения основной задачи. Обсудим предлагаемую схему.

Первая ее часть связана с проблемой выбора параметра регуляризации для экстремальных задач. Проблеме выбора параметра регуляризации α при решении некорректных задач посвящена обширная библиография (см., например, [3, 6, 8—20, 27—32, 41, 54—56, 59—65, 82, 86, 88, 96, 103—106, 110, 118—125, 130, 132—134, 141, 142, 149—152, 159—162, 164, 166, 181, 185—187, 190—206, 222—226, 230, 233]).

Способы выбора α обычно делятся на *априорные*, когда зависимость α от погрешности данных δ задачи задается заранее (см., например, [190—192]), и *апостериорные* (называемые иногда *принципами выбора* α), в которых величина α определяется не только по погрешностям, но и с использованием имеющихся приближенных реализаций данных задачи так, чтобы обеспечились определенные значения специальных функционалов от приближенного решения.

Для иллюстрации рассмотрим хорошо известный пример решения вариационным методом линейного операторного уравнения первого рода.

Пример 1. Пусть Z — гильбертово пространство с обычной топологией, A — линейный непрерывный оператор, действующий из Z в нормированное пространство U. Предположим, что уравнение

$$Az=\bar{u}, \quad z\in Z, \quad \bar{u}\in U, \tag{3}$$

имеет единственное решение $\bar{z}\in Z$. Тогда эта задача эквивалентна экстремальной задаче типа (2.2): найти $\bar{z}\in Z$ такое, что

$$\| A\bar{z}-\bar{u}\|_U=\inf\{\| Az-\bar{u}\|_U:\ z\in Z\}. \tag{4}$$

Если задать при этом $\Omega(z)=\|z\|_Z$ и считать, что вместо $\bar{u}$ известна приближенная правая часть $u_\delta\in U$ уравнения (3), подчиненная условию $\|u_\delta-\bar{u}\|_U\leqslant\delta$, то можно рассматривать задачу поиска приближений к $\bar{z}$ по данным $\{u_\delta, \delta\}$ как задачу приближенного определения Ω-оптимальных решений задачи (4). В этом случае можно использовать метод регуляризации, основанный на сглаживающем функционале вида

$$M_\delta^\alpha[z]=\alpha\|z\|_Z^2+\|Az-u_\delta\|_U^2, \quad z\in Z, \tag{5}$$

положив $f(x)=x^2$ $(x\geqslant 0)$ (см. [190, 191, 202]).

Задача (2) с $D=Z$ и со сглаживающим функционалом типа (5), как известно (см. § 0.5), имеет единственное решение $z_\delta^\alpha\in Z$ при любом $\alpha>0$. В [202] показано, что при выборе параметра $\alpha(\delta)$ из условий $\alpha(\delta)\to 0$, $\delta^2/\alpha(\delta)\to 0$ при $\delta\to 0$ элемент $z_\delta^{\alpha(\delta)}$ сходится в Z к $\bar{z}$. Такой выбор $\alpha(\delta)$ представляет собой пример априорного выбора параметра регуляризации. Возможен иной подход к выбору параметра α. В [86, 149] предложен так называемый *принцип невязки*, согласно которому при условии $\|u_\delta\|>\delta$ параметр α выбирается как корень уравнения

$$\beta(\alpha)\equiv\|Az_\delta^\alpha-u_\delta\|=\delta. \tag{6}$$

В этом случае значение $\alpha(\delta)$ определяется не только погрешностью δ, но и конкретным элементом u_δ (а также оператором A), входящим в условие выбора (6). Принцип невязки является примером апостериорного способа выбора параметра регуляризации. Он гарантирует не только сходимость приближений $z_\delta^{\alpha(\delta)}$ к $\bar{z}$ при

$\delta \to 0$, но и определенную величину невязки уравнения (3) на приближении $z_\delta^{\alpha(\delta)}$: $\| A z_\delta^{\alpha(\delta)} - u_\delta \| = \delta$.

Для вариационного метода решение линейных совместных операторных уравнений первого рода с точно заданным оператором A примерами регуляризирующих алгоритмов с апостериорным выбором параметра регуляризации служат также *принцип квазирешений* [86] и *принцип сглаживающего функционала* [150] (см. также [141, 142]). Эти алгоритмы хорошо зарекомендовали себя при решении большого числа практических задач [20, 56, 202, 204, 205, 207].

Нетрудно убедиться, однако, что упомянутые регуляризирующие алгоритмы не применимы для решения некорректных задач с возмущенными данными общего вида. Подробнее об этом сказано в § 2.2.

Регуляризирующие алгоритмы с априорным выбором параметра регуляризации для решений некорректных экстремальных задач достаточно хорошо изучены в работах [6, 8, 14, 27, 28, 30, 33, 88, 110, 142, 162, 164, 166, 179—182, 187, 190—199, 202, 203]. Поэтому в настоящей книге будут исследоваться в основном алгоритмы с апостериорным выбором α. Исключение составляет § 2.7.

В данной главе рассматриваются и обосновываются обобщенные аналоги принципов невязки, квазирешений и сглаживающего функционала для экстремальных задач. При этом оказывается, что не только выбор параметра регуляризации $\alpha(\delta)$ влияет в общем случае на сходимость приближенного решения $z_\delta \equiv z_\delta^{\alpha(\delta)}$ к множеству $\bar{Z}$ Ω-оптимальных решений. Оказывается, что неудачный выбор $z_\delta^{\alpha(\delta)}$ из множества $Z_\delta^{\alpha(\delta)}$ может привести к отсутствию искомой сходимости. Соответствующие примеры приведены в гл. 2 при рассмотрении решения нелинейных операторных уравнений вариационным методом. Таким образом, в общем случае необходимо в соответствии со второй частью схемы решения основной задачи производить отбор экстремали.

Заметим попутно, что в некоторых случаях в силу специфики выбора $\alpha(\delta)$ или же в силу особенностей рассматриваемой задачи (2.2) производить отбор экстремали нет необходимости. Примеры таких случаев приведены в § 4—7, 2.4—2.6. Такая же ситуация имеет место в примере 1.

В дальнейшем для краткости там, где речь не идет о сходимости при $\delta \to 0$, будем опускать индекс δ в символах $M_\delta^{\alpha(\delta)}$, Z_δ^α, $z_\delta^{\alpha(\delta)}$ и т. п., допуская соответственно запись $M^{\alpha(\delta)}$, Z^α, $z^{\alpha(\delta)}$.

В заключение параграфа сформулируем требования, которые будут предъявляться к вспомогательной функции $f(x)$, с помощью которой определяется сглаживающий функционал.

Определение 1. Будем говорить, что функция $f(x)$, заданная на $[x_0, +\infty)$, есть *сильно возрастающая функция порядка* m ($m \geqslant 1$), если выполнены условия: а) $f(x) \in C^m[x_0, +\infty)$; б) $f^{(r)}(x) \geqslant 0$ для $r = 1, \ldots, m-1$; в) $f^{(m)}(x) \geqslant \varkappa_0 = \mathrm{const} > 0$ для любого $x \in [x_0, +\infty)$.

Обозначим класс таких функций символом $\mathscr{F}^m = \mathscr{F}^m[x_0, +\infty)$ и в дальнейшем будем предполагать, что вспомогательная функция $f(x)$ принадлежит $\mathscr{F}^m$ для некоторого $m \geqslant 1$. Из формулы Тейлора следует, что сильно возрастающая функция порядка m $f(x)$ и ее производные $f^{(r)}(x)$ $(r=1, \ldots, m-1)$ — возрастающие на $[x_0, +\infty)$ функции. Это, в частности, означает, что $f(x) \in \mathscr{F}^m$ удовлетворяет условиям леммы 1 и теоремы 2.

При рассмотрении конкретных реализаций методики, предлагаемой в данном параграфе, часто используются в качестве вспомогательных функций $f(x)$ сильно возрастающие функции, составляющие класс $\mathscr{F}^1$, а также возрастающие и сильно выпуклые функции класса C^2, образующие класс $\mathscr{F}^2$. Простейшими функциями класса $\mathscr{F}^m$ являются $f(x)=x^m$ $(x \in [0, +\infty))$. Имеются функции, принадлежащие классам $\mathscr{F}^m$ для любого $m \geqslant 1$: $f(x)=e^x$ $(x \in [x_0, +\infty))$.

Отметим одно свойство функций класса $\mathscr{F}^m$ $(m>1)$.

Лемма 2. *Если $f(x) \in \mathscr{F}^m [x_0, +\infty)$ $(m>1)$, причем известно, что в некоторой точке x_1 $(x_1 \geqslant x_0)$ выполнено равенство $f'(x_1)=0$, то $x_1=x_0$, $f(x_1)=f(x_0)=\min\{f(x): x \geqslant x_0\}$.*

Доказательство. Предположим, что $x_1>x_0$. Тогда по формуле Тейлора

$$f(x_0)-f(x_1)=f'(x_1)(x_0-x_1)+f''(x_1+\theta(x_0-x_1))(x_0-x_1)^2/2=$$
$$=f''(x_1+\theta(x_0-x_1))(x_0-x_1)^2/2 \geqslant 0, \quad 0<\theta<1. \tag{7}$$

Здесь учтено, что для функции $f \in \mathscr{F}^m$ $(m>1)$ вторая производная не отрицательна (см. определение 1). Но вследствие возрастания $f(x)$ должно выполняться неравенство $f(x_0)<f(x_1)$, которое противоречит (7). Поэтому $x_1=x_0$, $f(x_1)=f(x_0)$. Лемма доказана.

§ 5. Некоторые следствия основных предположений

В данном параграфе рассматриваются вспомогательные утверждения, вытекающие из предположений относительно функционалов J, J_δ, Ω и функций Ψ, f, сделанных в § 2, 4. Эти утверждения будут использоваться в настоящей главе и гл. 2—4 при обосновании свойств изучаемых регуляризующих алгоритмов.

Большую роль при доказательстве сходимости регуляризующих алгоритмов будут играть следующие неравенства:

$$f[J^*-\Psi(\delta, \bar{\Omega})] \leqslant I_\delta(\bar{z}) \leqslant f[J^*+\Psi(\delta, \bar{\Omega})] \leqslant f[\lambda_\delta+\Psi(\delta, \bar{\Omega})], \tag{1}$$

$$M^\alpha[\bar{z}] \leqslant \alpha\bar{\Omega}+f[J^*+\Psi(\delta, \bar{\Omega})] \leqslant \alpha\bar{\Omega}+f[\lambda_\delta+\Psi(\delta, \bar{\Omega})], \tag{2}$$

справедливые для любого $\bar{z} \in \bar{Z}$. Чтобы убедиться в справедливости неравенства (1), достаточно использовать вытекающую из (2.3) оценку

$$J(\bar{z})-\Psi(\delta, \Omega(\bar{z})) \leqslant J_\delta(\bar{z}) \leqslant J(\bar{z})+\Psi(\delta, \Omega(\bar{z})) \quad \forall \bar{z} \in \bar{Z} \tag{3}$$

и принять во внимание равенства $J(\bar{z})=J^*$, $\Omega(\bar{z})=\bar{\Omega}$ (см. (2.4)), возрастание функции f, а также доказанное в теореме 3.1 неравенство $J^* \leqslant \lambda_\delta$. Неравенство (2) получается из (3) и определения сглаживающего функционала:

$$M^\alpha[\bar{z}]=\alpha\Omega(\bar{z})+f[J_\delta(\bar{z})] \leqslant \alpha\bar{\Omega}+f[J^*+\Psi(\delta, \bar{\Omega})] \leqslant \alpha\bar{\Omega}+f[\lambda_\delta+\Psi(\delta, \bar{\Omega})].$$

Нам понадобятся также некоторые утверждения, связанные с предельным переходом при $\Delta \equiv \|\delta\| \to 0$.

Лемма 1. *Предположим, что для функционалов J, J_δ, Ω выполнены предположения из § 2. Если для семейства элементов $\{z_\delta\}\subset D$ справедливы соотношения*

$$\lim_{\Delta\to 0} J_\delta(z_\delta)=J^*, \quad \overline{\lim_{\Delta\to 0}}\,\Omega(z_\delta)\leqslant\bar\Omega, \tag{4}$$

то для всякой последовательности $\{\delta_n\}$ такой, что $\Delta_n\equiv\|\delta_n\|\to 0$ при $n\to\infty$, соответствующая последовательность $\{z_n\}\equiv\{z_{\delta_n}\}$ τ-сходится к множеству $\bar Z$ Ω-оптимальных решений задачи (2.2) *при $n\to\infty$ и $\Omega(z_n)\to\bar\Omega$.*

Доказательство. Согласно второму из условий (4), семейство $\{z_\delta\}$ удовлетворяет требованию $\Omega(z_\delta)\leqslant\bar\Omega+\varepsilon_0\equiv K_0=\mathrm{const}$ при $\Delta\leqslant\Delta_0=\mathrm{const}$, т. е. $\{z_\delta\}\subset\Omega_{K_0}$. Поскольку вследствие преположения 2 из § 2 множество Ω_{K_0} τ-секвенциально компактно, то для любой последовательности $\{\delta_n\}$, сходящейся к нулю при $n\to\infty$, из соответствующей ей последовательности $\{z_{\delta_n}\}$ можно извлечь сходящуюся подпоследовательность $\{z_{\delta_{n_k}}\}\equiv\{z_{n_k}\}$: $z_{n_k}\xrightarrow{\tau}z^*\in D$. Из условия аппроксимации (2.3) и монотонности меры аппроксимации Ψ (см. предположение 3 из § 2) получим

$$J(z_{n_k})\leqslant J_{\delta_{n_k}}(z_{n_k})+\Psi(\delta_{n_k},\,\Omega(z_{n_k}))\leqslant J_{\delta_{n_k}}(z_{n_k})+\Psi(\delta_{n_k},\,K_0).$$

Переходя в этом неравенстве к пределу при $k\to\infty$ и учитывая τ-секвенциальную полунепрерывность снизу функционала J, предположения 3а), в) для функции Ψ, а также первое из условий (4), можно получить неравенство

$$J^*=\inf\{J(z):\ z\in D\}\leqslant J(z^*)\leqslant \varliminf_{k\to\infty} J(z_{n_k})\leqslant \lim_{k\to\infty}\,[J_{\delta_{n_k}}(z_{n_k})+\Psi(\delta_{n_k},\,K_0)]=$$

$$=\lim_{k\to\infty} J_{\delta_{n_k}}(z_{n_k})+\lim_{k\to\infty}\Psi(\delta_{n_k},\,K_0)=J^*.$$

Таким образом, z^* оказывается решением задачи (2.2). Используем теперь второе из условий (4). Оно вместе с (2.4) и τ-секвенциальной полунепрерывностью снизу функционала Ω (предположение 1 из § 2) дает

$$\bar\Omega=\inf\{\Omega(z):\ z\in Z^*\}\geqslant\overline{\lim_{\Delta\to 0}}\,\Omega(z_\delta)\geqslant\overline{\lim_{k\to\infty}}\,\Omega(z_{\delta_{n_k}})\geqslant\varliminf_{k\to\infty}\Omega(z_{n_k})\geqslant\Omega(z^*).$$

Из этого неравенства ясно, что z^* — Ω-оптимальное решение задачи (2.2), так что $\Omega(z^*)=\bar\Omega$ и $z_{n_k}\xrightarrow{\tau}z^*\in\bar Z$ при $k\to\infty$. Кроме того, из этого неравенства также следует, что $\Omega(z_{n_k})\to\bar\Omega$ при $k\to\infty$.

Докажем теперь, что вся последовательность $\{z_n\}$ τ-сходится к $\bar Z$. Пусть это не так, т. е. для некоторой окрестности $G\subset\tau$ множества $\bar Z$ и для некоторой подпоследовательности $\{n_p\}_{p=1}^{\infty}$ номеров оказывается, что $z_{n_p}\notin G$. При этом, согласно доказанному выше, из $\{z_{n_p}\}$ можно выбрать такую подпоследовательность $\{z_q\}\subset\{z_{n_p}\}$ $(q\equiv n_{p_l}$,

$l=1, 2, \ldots$), что $z_q \xrightarrow{\tau} \tilde{z} \in \tilde{Z}$, и, следовательно, по определению τ-сходимости $z_q \in G$ при $q=n_{p_l} \geqslant N(G)=\text{const}$. Полученное противоречие доказывает лемму.

Следствие 1. *Утверждения леммы* 1 *остаются справедливыми, если условие*

$$\lim_{\Delta \to 0} J_\delta(z_\delta)=J^*$$

заменить на условие

$$\lim_{\Delta \to 0} I_\delta(z_\delta)=\lim_{\Delta \to 0} f[J_\delta(z_\delta)]=f[J^*], \quad f(x) \in \mathcal{F}^m[x_0, +\infty).$$

Следствие 2. *Из условий и результата леммы* 1 *предположением от противного легко выводится, что* $\Omega(z_\delta) \to \tilde{\Omega}$ *при* $\|\delta\| \to 0$.

Лемма 2. *Если для семейства* $\{z_\delta\} \subset D$ *выполнены условия*

$$\lim_{\Delta \to 0} \Psi(\delta, \Omega(z_\delta))=0, \quad \overline{\lim_{\Delta \to 0}}\, I_\delta(z_\delta) \leqslant I^* \equiv f(J^*),$$

то $I_\delta(z_\delta) \to I^*$ *при* $\Delta \to 0$.

Доказательство. Из включения $f(x) \in \mathcal{F}^m \subset C^1[x_0, +\infty)$ ($m \geqslant 1$) вытекает справедливость формулы конечных приращений:

$$f(\lambda_\delta) \leqslant f[J_\delta(z_\delta)+\Psi(\delta, \Omega(z_\delta))]=$$
$$=f[J_\delta(z_\delta)]+f'[J_\delta(z_\delta)+\theta_1 \Psi(\delta, \Omega(z_\delta))]\Psi(\delta, \Omega(z_\delta)), \tag{5}$$
$$0<\theta_1<1, \quad \theta_1=\text{const}.$$

Учитывая здесь возрастание производной f' как следствие принадлежности $f(x) \in \mathcal{F}^m$, а также принимая во внимание вытекающие из условий леммы неравенства $\Psi(\delta, \Omega(z_\delta)) \leqslant \Psi_0=\text{const}$, $J_\delta(z_\delta) \leqslant J_0=\text{const}$, получаем

$$f(\lambda_\delta) \leqslant I_\delta(z_\delta)+f'[J_0+\Psi_0]\Psi(\delta, \Omega(z_\delta)).$$

Из этого неравенства, используя первое из условий леммы, доказанное в теореме 3.1 предельное соотношение $\lambda_\delta \to J^*$ при $\Delta \to 0$, а также учитывая непрерывность функций f, f', можно получить, что

$$I^*=f(J^*)=\lim_{\Delta \to 0}\{f(\lambda_\delta)-f'(J_0+\Psi_0)\Psi(\delta, \Omega(z_\delta))\} \leqslant \underline{\lim_{\Delta \to 0}}\, I_\delta(z_\delta).$$

Но тогда из второго условия леммы получается оценка

$$I^* \leqslant \underline{\lim_{\Delta \to 0}}\, I_\delta(z_\delta) \leqslant \overline{\lim_{\Delta \to 0}}\, I_\delta(z_\delta) \leqslant I^*,$$

которая и доказывает лемму.

Используя свойства функции Ψ, легко вывести из леммы 2

Следствие 3. *Если* $\Omega(z_\delta)\leqslant\Omega_0=\mathrm{const}$ *для любых* $\delta\in\mathbf{R}_+^k$ *таких, что* $\|\delta\|\leqslant\Delta_0$, *и если*

$$\overline{\lim_{\Delta\to 0}}\, I_\delta(z_\delta)\leqslant I^*,$$

то $I_\delta(z_\delta)\to I^*$ *при* $\Delta\to 0$.

Лемма 3. *Предположим, что функция* $\alpha(\delta)$ $(\delta\in\mathbf{R}_+^k)$ *удовлетворяет условию* $\alpha(\delta)\to 0$ *при* $\Delta\to 0$. *Далее будем считать, что для семейства* $\{z_\delta\}\subset D$ *выполнено неравенство*

$$M^{\alpha(\delta)}[z_\delta]\leqslant\alpha(\delta)\bar{\Omega}_\delta+f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]\quad\forall\delta,\ \|\delta\|\leqslant\Delta_0,$$

где числа $\bar{\Omega}_\delta$ *равномерно ограничены*: $\bar{\Omega}_\delta\leqslant\tilde{\Omega}_0=\mathrm{const}$ *при* $\|\delta\|\leqslant\Delta_0$. *Тогда для* z_δ *справедливо соотношение*

$$\overline{\lim_{\Delta\to 0}}\, I_\delta(z_\delta)\leqslant I^*. \tag{6}$$

Доказательство. Из условия леммы и определения сглаживающего функционала вытекает оценка

$$I_\delta(z_\delta)=M^{\alpha(\delta)}[z_\delta]-\alpha(\delta)\Omega(z_\delta)\leqslant\alpha(\delta)[\bar{\Omega}_\delta-\Omega^*]+$$
$$+f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]\leqslant\alpha(\delta)[\tilde{\Omega}_0-\Omega^*]+f[\lambda_\delta+\Psi(\delta,\tilde{\Omega}_0)].$$

Правая часть этой оценки на основании свойств функций f, Ψ, а также с учетом сходимостей $\alpha(\delta)\to 0$, $\lambda_\delta\to J^*$ при $\Delta\to 0$ стремится к $f(J^*)$, что и требуется доказать.

Следствие 4. *Пусть* $z_\delta\equiv z^{\alpha(\delta)}$ — *экстремаль сглаживающего функционала в задаче* (4.2) *с* $\alpha=\alpha(\delta)$. *Тогда, если* $\alpha(\delta)\to 0$ *при* $\Delta\to 0$, *то для* z_δ *справедливо* (6).

Для проверки этого достаточно положить $\bar{\Omega}_\delta\equiv\bar{\Omega}$ и убедиться в том, что выполнено условие леммы 3, так как, согласно (2), и по определению экстремали

$$M^{\alpha(\delta)}[z^{\alpha(\delta)}]\leqslant M^{\alpha(\delta)}[\bar{z}]\leqslant\alpha(\delta)\bar{\Omega}+f[\lambda_\delta+\Psi(\delta,\bar{\Omega})]\quad\forall\bar{z}\in\bar{Z}.$$

Из лемм 2, 3 и следствия 3 вытекает также

Следствие 5. *Пусть* $\alpha(\delta)\to 0$ *и семейство* $\{z_\delta\}\subset D$ *удовлетворяет условиям леммы 3. Тогда при выполнении одного из условий*: а) $\Psi(\delta,\Omega(z_\delta))\to 0$ *при* $\Delta\to 0$; б) $\Omega(z_\delta)\leqslant\Omega_0$ *при* $\Delta\leqslant\Delta_0$, *имеет место сходимость* $I_\delta(z_\delta)\to I^*$ *при* $\Delta\to 0$. *В частности, если* $z^{\alpha(\delta)}$ — *экстремаль задачи* (4.2) *для* $\alpha=\alpha(\delta)$ *и* $\Psi(\delta,\Omega(z_\delta^{\alpha(\delta)}))\to 0$ *или* $\Omega(z_\delta^{\alpha(\delta)})\leqslant\Omega_0$, *то* $I_\delta(z_\delta^{\alpha(\delta)})\to I^*$.

Лемма 4. *Предположим, что для элементов* $z_\delta\in D$ *и функции* $\alpha(\delta)>0$ *выполнено условие*

$$M^{\alpha(\delta)}[z_\delta]\leqslant\alpha(\delta)\bar{\Omega}_\delta+f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]\quad\forall\delta\in\mathbf{R}_+^k,\quad\|\delta\|\leqslant\Delta_0.$$

Тогда справедливо неравенство

$$\Omega(z_\delta)-\bar{\Omega}_\delta\leqslant\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-I_\delta(z_\delta)\}/\alpha(\delta)\leqslant$$
$$\leqslant\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-f_{\min}\}/\alpha(\delta), \tag{7}$$

где $f_{\min}\equiv\min\{f(x)\colon\ x\geqslant x_0\}=f(x_0)$. *В частности, если положить* $z_\delta=z^{\alpha(\delta)}\in Z^{\alpha(\delta)}$ *и считать, что* $\bar{\Omega}_\delta\geqslant\bar{\Omega}$, *то*

$$\Omega(z^{\alpha(\delta)})-\bar{\Omega}_\delta\leqslant[I_\delta(\bar{z})-I_\delta(z^{\alpha(\delta)})]/\alpha(\delta)\leqslant$$
$$\leqslant\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-I_\delta(z^{\alpha(\delta)})\}/\alpha(\delta)\leqslant\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-f_{\min}\}/\alpha(\delta). \quad (7^*)$$

Доказательство. Формула (7) получается непосредственно из условия леммы и из следствия равенства (4.1):

$$\Omega(z_\delta)-\bar{\Omega}_\delta=[M^{\alpha(\delta)}[z_\delta]-\alpha(\delta)\bar{\Omega}_\delta-I_\delta(z_\delta)]/\alpha(\delta).$$

Чтобы обосновать (7*), достаточно заметить, что из (2), (4.2)

$$M^{\alpha(\delta)}[z^{\alpha(\delta)}]\leqslant M^{\alpha(\delta)}[\bar{z}]\leqslant\alpha(\delta)\bar{\Omega}+f[J^*+\Psi(\delta,\bar{\Omega})]\leqslant$$
$$\leqslant\alpha(\delta)\bar{\Omega}_\delta+f[J^*+\Psi(\delta,\bar{\Omega}_\delta)].$$

Здесь учтена также монотонность функций f, Ψ. Таким образом, для $z_\delta=z^{\alpha(\delta)}$ выполнено условие леммы 4, из которой и следует (7*) как частный случай (7).

Лемма 5. *Пусть элементы* $z_\delta\in D$ *и функция* $\alpha(\delta)$ *удовлетворяют условиям леммы* 4, *причем* $\alpha(\delta)\to 0$ *при* $\Delta\to 0$. *Пусть, кроме того, фигурирующие в лемме* 4 *числа* $\bar{\Omega}_\delta$ *подчинены условию* $\bar{\Omega}_\delta\to\bar{\Omega}$ *при* $\Delta\to 0$. *Если при этом выполнены соотношения*

$$\overline{\lim_{\Delta\to 0}}\, I_\delta(z_\delta)\leqslant I^*, \quad (8)$$

$$\lim_{\Delta\to 0}\zeta_1(\delta)\equiv\lim_{\Delta\to 0}[\Psi(\delta,\Omega(z_\delta))/\alpha(\delta)]=0, \quad (9)$$

$$\lim_{\Delta\to 0}\zeta_2(\delta)\equiv\lim_{\Delta\to 0}\{f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]-f(\lambda_\delta)\}/\alpha(\delta)=0, \quad (10)$$

то

$$\overline{\lim_{\Delta\to 0}}\,\Omega(z_\delta)\leqslant\bar{\Omega},\quad \overline{\lim_{\Delta\to 0}}\, I_\delta(z_\delta)=I^*.$$

Доказательство. Из формул (5), (7), а также из неравенства $J^*\leqslant\lambda_\delta$ (см. теорему 3.1) получим оценку

$$\Omega(z_\delta)-\bar{\Omega}_\delta\leqslant\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-f[J_\delta(z_\delta)]\}/\alpha(\delta)\leqslant\{f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]-$$
$$-f(\lambda_\delta)+f'[J_\delta(z_\delta)+\Psi(\delta,\Omega(z_\delta))]\,\Psi(\delta,\Omega(z_\delta))\}/\alpha(\delta). \quad (11)$$

Величины $\Psi(\delta,\Omega(z_\delta))$ ограничены в совокупности, в чем можно убедиться из (9) и из сходимости $\alpha(\delta)\to 0$ при $\Delta\to 0$. Величины $J_\delta(z_\delta)$ также равномерно по δ ограничены сверху, что вытекает из (8). Поэтому из (11) с учетом монотонности функции f' можно получить

$$\Omega(z_\delta)-\bar{\Omega}_\delta\leqslant\zeta_2(\delta)+f'(Q_0)\zeta_1(\delta),\quad Q_0=\text{const}.$$

Из этого неравенства и из (9), (10) следует первое доказываемое предельное соотношение. Оно вместе с условием (8) доказывает, согласно следствию 3, и второе требуемое соотношение.

Лемма 6. *Предположим, что* $\bar{\Omega}_\delta\to\bar{\Omega}$ *при* $\Delta\to 0$. *Если для семейства* $\{z_\delta\}\subset D$ *справедливы оценки*

$$M^{\alpha_0}[z_\delta]\leqslant\alpha_0\bar{\Omega}_\delta+f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]\quad\forall\delta\in\mathrm{R}^k_+,\quad \|\delta\|\leqslant\Delta_0.$$

где α_0 — не зависящая от δ положительная константа, то величины $\Omega(z_\delta)$ равномерно ограничены по δ: $\Omega(z_\delta)\leqslant\tilde{\Omega}=\text{const}$. *В частности, если $z_\delta=z_\delta^{\alpha_0}\in Z_\delta^{\alpha_0}$ — экстремаль задачи* (4.2) *с* $\alpha=\alpha_0$, *то так же* $\Omega(z_\delta^{\alpha_0})\leqslant\tilde{\Omega}$ *для любого* δ $(\|\delta\|\leqslant\Delta_0)$.

Доказательство. Условия леммы 6 обеспечивают выполнение условия леммы 4 с $\alpha(\delta)=\alpha_0>0$. Тогда по лемме 4 (см. (7) или (7*))

$$\Omega(z_\delta)\leqslant\bar{\Omega}_\delta+\{f[J^*+\Psi(\delta,\bar{\Omega}_\delta)]-f_{\min}\}/\alpha_0. \tag{12}$$

Поскольку величины $\bar{\Omega}_\delta$ ограничены вследствие сходимости $\bar{\Omega}_\delta\to\bar{\Omega}$ $(\Delta\to 0)$, то выполнена оценка $0\leqslant\Psi(\delta,\bar{\Omega}_\delta)\leqslant\Psi(\delta,\Omega_1)$, где $\Omega_1=\text{const}\geqslant\Omega^*$. Эта оценка в силу свойств функции Ψ обеспечивает сходимости $\Psi(\delta,\Omega_1)\to 0$ и $\Psi(\delta,\bar{\Omega}_\delta)\to 0$ при $\Delta\to 0$. Поэтому величины $\Psi(\delta,\bar{\Omega}_\delta)$ ограничены: $\Psi(\delta,\bar{\Omega}_\delta)\leqslant\bar{\Psi}=\text{const}$ для любого δ $(\|\delta\|\leqslant\Delta_0)$. Тогда из (12) получим

$$\Omega(z_\delta)\leqslant\Omega_1+\{f[J^*+\bar{\Psi}]-f_{\min}\}/\alpha_0\equiv\tilde{\Omega}=\text{const},$$

что и требовалось доказать.

Предположим теперь, что имеется семейство элементов $z_\delta\equiv z[\alpha(\delta)]\in D$, зависящих от величины $\alpha(\delta)>0$, а также заданы два вспомогательных семейства, связанных с z_δ:

$$z_\delta^{(1)}\equiv z^{(1)}[\alpha(\delta)]\in D,\quad z_\delta^{(2)}\equiv z^{(2)}[\alpha(\delta)]\in D.$$

Будем считать далее, что при всех $\delta\in\mathbf{R}_+^k$ $(\|\delta\|\leqslant\Delta_0)$ эти семейства удовлетворяют следующим условиям:

а) выполнено неравенство для z_δ

$$M_\delta^{\alpha(\delta)}[z_\delta]\equiv\alpha(\delta)\Omega(z_\delta)+I_\delta(z_\delta)\leqslant\alpha(\delta)\bar{\Omega}_\delta+f[J^*+\Psi(\delta,\bar{\Omega}_\delta)],$$

где числа $\bar{\Omega}_\delta$ стремятся к $\bar{\Omega}$ при $\Delta\to 0$;

б) для семейства $z_\delta^{(2)}$ выполняется неравенство

$$M_\delta^{\alpha(\delta)q}[z_\delta^{(2)}]\leqslant\alpha(\delta)q\bar{\Omega}_\delta+f[J^*+\Psi(\delta,\bar{\Omega}_\delta)],$$

где снова $\bar{\Omega}_\delta\to\bar{\Omega}$ при $\Delta\to 0$; q — произвольная константа, большая единицы;

в) семейства $z_\delta^{(1,2)}$ и z_δ связаны между собой неравенствами

$$\Omega(z_\delta^{(1)})\geqslant\Omega(z_\delta)\geqslant\Omega(z_\delta^{(2)}),\quad I_\delta(z_\delta^{(1)})\leqslant I_\delta(z_\delta)\leqslant I_\delta(z_\delta^{(2)});$$

г) семейства $z_\delta^{(1,2)}$ связаны с z_δ так, что из выполнения неравенства

$$I_\delta(z_\delta^{(2)})\geqslant Cf[\lambda_\delta+\Psi(\delta,\Omega(z_\delta^{(1)}))]-(C-1)f(\lambda_\delta) \tag{13}$$

следует, что

$$I_\delta(z_\delta)\leqslant f[\lambda_\delta+\Psi(\delta,\Omega(z_\delta))], \tag{14}$$

а из выполнения неравенства

$$I_\delta(z_\delta^{(2)})\leqslant Cf[\lambda_\delta+\Psi(\delta,\Omega(z_\delta^{(1)}))]-(C-1)f(\lambda_\delta) \tag{15}$$

следует, что

$$I_\delta(z_\delta)\geqslant f[\lambda_\delta+\Psi(\delta,\Omega(z_\delta))]; \tag{16}$$

здесь $C>1$ — константа, не зависящая от δ;

д) если для любого $\delta\in\mathbf{R}^k_+$ ($\|\delta\|\leqslant\Delta_0$) величины $\alpha(\delta)$ подчинены требованию $\alpha(\delta)\geqslant\alpha_0=\mathrm{const}>0$, то выполнено неравенство $\Omega(z^{(1)}_\delta)\leqslant\leqslant\Omega(z^{\alpha_0/q}_\delta)$, где $q=\mathrm{const}>1$, а $z^{\alpha_0/q}_\delta$ — экстремаль задачи (4.2) с $\alpha=\alpha_0/q$.

При этих условиях семейство $\{z_\delta\}$ обладает свойствами, указанными в следующих леммах.

Лемма 7. *Если выполнены условия* а) — д), *причем конкретно из* (15) *следует* (16), *то* $\Omega(z_\delta)\leqslant\bar{\Omega}_\delta$. *Если же* $\Omega(z_\delta)>\bar{\Omega}_\delta$, *то выполнены* (13), (14), *причем неравенство* (14) *строгое*.

Доказательство. В силу условия а) из леммы 4 следует

$$\Omega(z_\delta)-\bar{\Omega}_\delta\leqslant\{f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]-I_\delta[z_\delta]\}/\alpha(\delta)\leqslant$$
$$\leqslant\{f[\lambda_\delta+\Psi(\delta,\bar{\Omega}_\delta)]-f[\lambda_\delta+\Psi(\delta,\Omega(z_\delta))]\}/\alpha(\delta),$$

где использовано (16) и неравенство $\lambda_\delta\geqslant J^*$ (см. теорему 3.1). Отсюда и из монотонности функций f, Ψ следует, что $\Omega(z_\delta)\leqslant\bar{\Omega}_\delta$, и это доказывает первую часть леммы. Справедливость второй части леммы непосредственно вытекает из первой части.

Лемма 8. *Если для любого* $\delta\in\mathbf{R}^k_+$ ($\|\delta\|\leqslant\Delta_0$) *выполнено неравенство* $\alpha(\delta)\geqslant\alpha_0>0$ ($\alpha_0=\mathrm{const}$), *то* $I_\delta(z_\delta)\to I^*$ *при* $\Delta\to0$.

Доказательство. По условию д) с учетом в) имеем

$$\Omega(z_\delta)\leqslant\Omega(z^{(1)}_\delta)\leqslant\Omega(z^{\alpha_0/q}_\delta)\leqslant\tilde{\Omega}_0=\mathrm{const}, \tag{17}$$

где последнее неравенство получено в соответствии с леммой 6. Возможны два случая. В первом выполнены (13) и, следовательно, (14). Тогда из (14), (17) с учетом монотонности функций f, Ψ получим

$$I_\delta(z_\delta)\leqslant f[\lambda_\delta+\Psi(\delta,\Omega(z_\delta))]\leqslant f[\lambda_\delta+\Psi(\delta,\tilde{\Omega}_0)]\to I^*, \quad \Delta\to0. \tag{18}$$

Если же выполнены (15), (16), то на основании неравенства (15), принимая во внимание монотонность функций f, Ψ, неравенство (17), условие в), а также доказанную в теореме 3.1 сходимость $\lambda_\delta\to J^*$ при $\Delta\to0$, можно сделать вывод, что

$$I_\delta(z_\delta)\leqslant I_\delta(z^{(2)}_\delta)\leqslant Cf[\lambda_\delta+\Psi(\delta,\Omega(z^{(1)}_\delta))]-(C-1)f(\lambda_\delta)\leqslant$$
$$\leqslant C\{f[\lambda_\delta+\Psi(\delta,\tilde{\Omega}_0)]-f(\lambda_\delta)\}+f(\lambda_\delta)\to f(J^*),\quad \Delta\to0. \tag{19}$$

Из неравенств (18), (19) ясно, что $\overline{\lim\limits_{\Delta\to0}}\, I_\delta(z_\delta)\leqslant I^*$. Тогда по следствию 3 с учетом (17) получим доказываемое утверждение.

Лемма 9. *Если* $\alpha(\delta)\to0$ *при* $\Delta\to0$, *то* $I_\delta(z_\delta)\to I^*$.

Доказательство. Рассмотрим два случая. В первом для каждого δ имеем $\Omega(z_\delta)\leqslant\bar{\Omega}_\delta$. Тогда в силу ограниченности величин $\bar{\Omega}_\delta$ справедливо следствие 5, откуда вытекает справедливость утверждения леммы в рассматриваемом случае.

Во втором случае

$$\Omega(z_\delta)>\bar{\Omega}_\delta. \tag{20}$$

Тогда по лемме 7 выполнены (13), (14), и из (13) с учетом (20) и условия в) получим

$$I_\delta(z_\delta^{(2)}) \geqslant Cf[\Psi(\delta, \Omega(z_\delta^{(1)})) + \lambda_\delta] - (C-1) f(\lambda_\delta) \geqslant$$
$$\geqslant Cf[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta] - (C-1) f(\lambda_\delta), \quad (21)$$

где использована также монотонность функций f, Ψ. Из (21) и условия б) следует

$$\alpha(\delta) q \Omega(z_\delta^{(2)}) + Cf[\Psi(\delta, \Omega(z_\delta^{(1)})) + \lambda_\delta] - (C-1) f(\lambda_\delta) \leqslant$$
$$\leqslant \alpha(\delta) q \Omega(z_\delta^{(2)}) + I_\delta(z_\delta^{(2)}) = M_\delta^{\alpha(\delta) q}[z_\delta^{(2)}] \leqslant \alpha(\delta) q \bar{\Omega}_\delta + f[\lambda_\delta + \Psi(\delta, \bar{\Omega}_\delta)]. \quad (22)$$

Неравенства (21), (22) дают

$$0 \leqslant (C-1)\{f[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta] - f(\lambda_\delta)\}/\alpha(\delta) \leqslant$$
$$\leqslant C\{f[\Psi(\delta, \Omega(z_\delta^{(1)})) + \lambda_\delta] - f(\lambda_\delta)\}/\alpha(\delta) +$$
$$+ \{f(\lambda_\delta) - f[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta]\}/\alpha(\delta) \leqslant q[\bar{\Omega}_\delta - \Omega(z_\delta^{(2)})]. \quad (23)$$

Отсюда получаем

$$\Omega(z_\delta^{(2)}) \leqslant \bar{\Omega}_\delta \to \bar{\Omega}, \quad \Delta \to 0. \quad (24)$$

Поскольку $\alpha(\delta) \to 0$, то по следствию 5 неравенство (24) обеспечивает сходимость $I_\delta(z_\delta^{(2)}) \to I^*$ $(\Delta \to 0)$. Из этой сходимости вместе с (24) на основании следствия 2 следует, что $\Omega(z_\delta^{(2)}) \to \bar{\Omega}$ $(\Delta \to 0)$. Это в свою очередь приводит к следующим следствиям из (23):

$$\{f[\lambda_\delta + \Psi(\delta, \bar{\Omega}_\delta)] - f(\lambda_\delta)\}/\alpha(\delta) \to 0, \quad (25)$$
$$\omega(\delta) \equiv \{f[\lambda_\delta + \Psi(\delta, \Omega(z_\delta^{(1)}))] - f(\lambda_\delta)\}/\alpha(\delta) \to 0, \quad (26)$$
$$\Delta \to 0.$$

Из (26) с учетом непрерывности f и сходимости $\alpha(\delta) \to 0$ ясно, что $\Psi(\delta, \Omega(z_\delta^{(1)})) \to 0$ при $\Delta \to 0$. Поскольку $0 \leqslant \Psi(\delta, \Omega(z_\delta)) \leqslant \Psi(\delta, \Omega(z_\delta^{(1)}))$, то и $\Psi(\delta, \Omega(z_\delta)) \to 0$ при $\Delta \to 0$. Но тогда для z_δ выполнены условия следствия 5, согласно которому лемма 9 оказывается верной и во втором случае.

В процессе доказательства леммы 9 фактически установлено

Следствие 6. *При выполнении условий* а)—д) *из условий* $\alpha(\delta) \to 0$ *при* $\Delta \to 0$, $\Omega(z_\delta) > \bar{\Omega}_\delta$ *при* $\|\delta\| \leqslant \Delta_0$ *следуют соотношения* (25), (26).

Из лемм 8, 9 можно получить

Следствие 7. *При выполнении условий* а)—д) $I_\delta(z_\delta) \underrightarrow{\ } I^*$ *при* $\Delta \to 0$.

Лемма 10. *Если выполнены условия* а)—д), *то* $\varlimsup_{\Delta \to 0} \Omega(z_\delta) \leqslant \bar{\Omega}$.

Доказательство. Рассмотрим два случая. В первом случае $\Omega(z_\delta) \leqslant \bar{\Omega}_\delta$ для любого $\delta \in \mathbf{R}_+^k$ $(\|\delta\| \leqslant \Delta_0)$, и лемма очевидна.

Во втором случае выполнено неравенство

$$\Omega(z_\delta) > \bar{\Omega}_\delta. \quad (27)$$

Здесь необходимо изучить два варианта. Первый из них предусматривает выполнение для любого δ $(\|\delta\| \leqslant \Delta_0)$ неравенства $\alpha(\delta) \geqslant \alpha_0 > 0$. Тогда (27) вместе с условием а) на основании леммы 4 дают

$$0 \leqslant \Omega(z_\delta) - \bar{\Omega}_\delta \leqslant \{f[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta] - I_\delta(z_\delta)\}/\alpha_0. \quad (28)$$

Поскольку по условию а) $\bar{\Omega}_\delta \to \bar{\Omega}$ при $\Delta \to 0$, то и $\Psi(\delta, \bar{\Omega}_\delta) \to 0$, как отмечалось уже при доказательстве леммы 6. Но тогда из (28) и из следствия 7 получается сходимость $\Omega(z_\delta) \to \bar{\Omega}$ при $\Delta \to 0$.

Во втором, более сложном варианте предполагается, что $\alpha(\delta) \to 0$ при $\Delta \to 0$. Тогда на основании неравенства (27) можно утверждать, что выполнены условия следствия 6, и, значит, справедливы предельные соотношения (25), (26). Дальнейшая процедура доказательства зависит от того, какому классу $\mathcal{F}^m$ принадлежит вспомогательная функция $f(x)$.

Так, например, если $f(x) \in \mathcal{F}^1$, то из (26) по формуле конечных приращений получаем, принимая во внимание условие в) и монотонность функции Ψ,

$$\omega(\delta) = f'(\lambda_\delta^*)\Psi(\delta, \Omega(z_\delta^{(1)}))/\alpha(\delta) \geqslant \varkappa_0 \Psi(\delta, \Omega(z_\delta))/\alpha(\delta) \geqslant 0$$

(см. также определение 4.1). Здесь λ_δ^* — некоторое число, удовлетворяющее неравенству $\lambda_\delta^* > \lambda_\delta$. Поскольку в силу (26) $\omega(\delta) \to 0$ при $\Delta \to 0$, то и $\Psi(\delta, \Omega(z_\delta))/\alpha(\delta) \to 0$, если $\Delta \to 0$. Этот факт вместе со справедливым по-прежнему соотношением (25), а также с установленной в следствии 7 сходимостью $I_\delta(z_\delta) \to I^*$ при $\Delta \to 0$ представляет собой условия леммы 5. Из нее и получается требуемое неравенство

$$\varlimsup_{\Delta \to 0} \Omega(z_\delta) \leqslant \bar{\Omega}.$$

Аналогичные рассуждения допустимы и для $f(x) \in \mathcal{F}^m$ $(m > 1)$, если, кроме того, выполняется неравенство $f'(J^*) > 0$. Действительно, в этом случае, как и при $f \in \mathcal{F}^1$,

$$\omega(\delta) = f'(\lambda_\delta^*)\Psi(\delta, \Omega(z_\delta^{(1)}))/\alpha(\delta) \geqslant f'(J^*)\Psi(\delta, \Omega(z_\delta))/\alpha(\delta) \geqslant 0,$$

где учтены вытекающее из теоремы 3.1 неравенство $\lambda_\delta^* > \lambda_\delta \geqslant J^*$ и монотонность производной f'. Все остальные выводы остаются без изменения.

Остается рассмотреть такие функции $f(x)$, для которых $f \in \mathcal{F}^m$ $(m > 1)$, но $f'(J^*) = 0$. Как установлено в лемме 4.2, для таких $f(x)$ оказывается, что $f(J^*) = f_{\min}$. Из условия а) и леммы 8 тогда вытекает (7) в следующей форме:

$$\Omega(z_\delta) - \bar{\Omega}_\delta \leqslant \{f[J^* + \Psi(\delta, \bar{\Omega}_\delta)] - f(J^*)\}/\alpha(\delta).$$

Используя (27) и раскладывая правую часть этого неравенства по формуле Тейлора, можно проделать следующую выкладку:

$$0 < \Omega(z_\delta) - \bar{\Omega}_\delta \leqslant \frac{1}{\alpha(\delta)}\{f[J^* + \Psi(\delta, \bar{\Omega}_\delta) - f(J^*)\} =$$

$$= \sum_{k=1}^{m-1} \frac{f^{(k)}(J^*)\Psi^k(\delta, \bar{\Omega}_\delta)}{k!\ \alpha(\delta)} + \frac{1}{m!} f^{(m)}[J^* + \theta\Psi(\delta, \bar{\Omega}_\delta)] \frac{\Psi^m(\delta, \bar{\Omega}_\delta)}{\alpha(\delta)} \leqslant$$

$$\leqslant \sum_{k=1}^{m-1} \frac{f^{(k)}(J^*)}{k!} \frac{\Psi^k(\delta, \Omega(z_\delta))}{\alpha(\delta)} + \frac{1}{m!} \max_{[J^*, J_1]} \{f^{(m)}(x)\} \frac{\Psi^m(\delta, \Omega(z_\delta))}{\alpha(\delta)} \leqslant$$

$$\leqslant \sum_{k=1}^{m-1} \frac{f^{(k)}(\lambda_\delta)}{k!} \frac{\Psi^k(\delta, \Omega(z_\delta^{(1)}))}{\alpha(\delta)} + N_1 \frac{\Psi^m(\delta, \Omega(z_\delta^{(1)}))}{\alpha(\delta)} \equiv \Delta_1(\delta) + N_1\Delta_2(\delta) \quad (29)$$

Здесь θ $(0<\theta<1)$ — константа остаточного члена формулы Тейлора, J_1 — константа, ограничивающая сверху множество значений величины $J^*+\Psi(\delta, \bar{\Omega}_\delta)$ при $\|\delta\|\leqslant\Delta_0$. Константа

$$N_1\equiv\frac{1}{m!}\max_{[J^*, J_1]} f^{(m)}(x)$$

конечна в силу непрерывности производной $f^{(m)}$. В выкладке (29) использованы монотонное возрастание производных $f^{(k)}$ $(1\leqslant k\leqslant m)$, отмеченное в § 4, неравенство $J^*\leqslant\lambda_\delta$, а также вытекающее из условия в) и из (27) неравенство $\bar{\Omega}_\delta<\Omega(z_\delta)\leqslant\Omega(z_\delta^{(1)})$.

Изучим теперь поведение величин $\Delta_{1,2}(\delta)$ при $\|\delta\|\to 0$. Для этого напишем снова формулу Тейлора для функции $\omega(\delta)$ из (26) и проведем следующую оценку:

$$\omega(\delta)=\frac{1}{\alpha(\delta)}\{f[\lambda_\delta+\Psi(\delta, \Omega(z_\delta^{(1)}))]-f(\lambda_\delta)\}=$$

$$=\sum_{k=1}^{m-1}\frac{f^{(k)}(\lambda_\delta)}{k!}\,\frac{\Psi^k(\delta, \Omega(z_\delta^{(1)}))}{\alpha(\delta)}+\frac{1}{m!}f^{(m)}[\lambda_\delta+\theta_1\Psi(\delta, \Omega(z_\delta^{(1)}))]\frac{\Psi^m(\delta, \Omega(z_\delta^{(1)}))}{\alpha(\delta)}\geqslant$$

$$\geqslant\Delta_1(\delta)+\frac{\varkappa_0}{m!}\Delta_2(\delta). \qquad (30)$$

Здесь $\theta_1=\text{const}$ $(0<\theta_1<1)$ и принято во внимание, что по определению 4.1 для функции $f\in\mathscr{F}^m$ верна оценка $f^{(m)}(x)\geqslant\varkappa_0$ для любого $x\geqslant x_0$.

Как установлено в лемме 9, $\omega(\delta)\to 0$ при $\|\delta\|\to 0$ (см. (26)). Поэтому из оценки (30) можно сделать вывод, что неотрицательные величины $\Delta_{1,2}(\delta)$ сходятся к нулю при $\Delta\to 0$. Но тогда предельный переход при $\Delta\to 0$ в неравенстве (29) дает сходимость $\Omega(z_\delta)\to\bar{\Omega}$, так как $\bar{\Omega}_\delta\to\bar{\Omega}$ при $\Delta\to 0$. Лемма доказана.

Итак, условия а) — д), согласно следствию 7 и лемме 10, гарантируют выполнение двух предельных соотношений для семейства $\{z_\delta\}$:

$$\lim_{\Delta\to 0} I_\delta(z_\delta)=I^*, \qquad (31)$$

$$\overline{\lim_{\Delta\to 0}}\,\Omega(z_\delta)\leqslant\bar{\Omega}. \qquad (32)$$

Лемма 11. *Пусть в условиях* а) — д) *вместо неравенств* (14), (16) *используются соответственно неравенства* $\Omega(z_\delta)\geqslant\bar{\Omega}_\delta$, $\Omega(z_\delta)\leqslant\bar{\Omega}_\delta$. *Тогда снова справедливы формулы* (31), (32).

Доказательство. Установим (31). Для этого сначала рассмотрим случай $\alpha(\delta)\geqslant\alpha_0>0$, разбираемый в лемме 8 для случая, когда выполнены (14), (16). Все приведенное там доказательство сохраняется и при выполнении условий леммы 11, за исключением неравенства (18), которое там выводится из (14).

В нашем случае вместо (18) необходимо использовать вытекающее из условия а) и из условия $\Omega(z_\delta)\geqslant\bar{\Omega}_\delta$ леммы неравенство

$$I_\delta(z_\delta)=M_\delta^{\alpha(\delta)}[z_\delta]-\alpha(\delta)\Omega(z_\delta)\leqslant\alpha(\delta)[\bar{\Omega}_\delta-\bar{\Omega}_\delta]+$$
$$+f[J^*+\Psi(\delta, \bar{\Omega}_\delta)]=f[J^*+\Psi(\delta, \bar{\Omega}_\delta)]\to f(J^*)=I^*, \quad \Delta\to 0.$$

Из этого неравенства получается тот же вывод, что и в лемме 8 из (19), так что дальнейшие рассуждения по схеме из леммы 8 дают (31). Доказательство случая $\alpha(\delta)\to 0$ при $\Delta\to 0$, разбираемого в лемме 9, остается без изменений.

Доказательство неравенства (32) проводится в точном соответствии со схемой, приведенной в лемме 10, и использует уже доказанное равенство (31).

При исследовании свойств регуляризующих алгоритмов будут встречаться две вспомогательные экстремальные задачи. Первая: найти элементы $z_0\in D$, для которых

$$\Omega(z_0)=\inf\{\Omega(z)\colon z\in D\}\equiv\Omega^*. \tag{33}$$

Будем обозначать множество ее решений как $Z_0\equiv\{z\in D\colon \Omega(z)=\Omega^*\}$.

Тогда вторая задача формулируется следующим образом: найти элементы $\hat{z}_\delta\in Z_0$ такие, что

$$I_\delta(\hat{z}_\delta)=\inf\{I_\delta(z)\colon z\in Z_0\}\equiv\nu_\delta. \tag{34}$$

Существование решений этих задач обосновывает

Лемма 12. *Пусть для* J_δ, Ω *выполнены предположения* 1, 2 *из* § 2, *а* $f(x)$ *непрерывна и монотонно возрастает* (*в частности, может быть* $f(x)\in\mathcal{F}^m$ ($m\geqslant 1$)). *Тогда задачи* (33), (34) *разрешимы. Множества их решений* Z_0, Z_0^δ *обладают свойством* τ*-секвенциальной компактности.*

Доказательство проводится по той же схеме, что и в теореме 4.2, и поэтому опускается.

§ 6. Вспомогательные функции и их свойства

Предположим, что задача (4.2) минимизации сглаживающего функционала: найти

$$Z_\delta^\alpha\equiv\operatorname{Arg\,inf}\{M_\delta^\alpha[z]\colon z\in D\}$$

имеет непустое множество решений Z_δ^α для всякого $\alpha>0$ при любом фиксированном $\delta\in\mathbf{R}_+^k$ ($\|\delta\|\leqslant\Delta_0$). В частности, это будет иметь место при выполнении предположений из § 2, 4 о функционалах J_δ, Ω и функции f (см. теорему 4.2). Тогда можно ввести многозначные функции аргумента $\alpha>0$

$$\gamma(\alpha)\equiv\Omega(z^\alpha),\quad \beta(\alpha)\equiv I_\delta(z^\alpha),\quad \varphi(\alpha)\equiv M_\delta^\alpha[z^\alpha]\quad \forall z^\alpha\in Z_\delta^\alpha \tag{1}$$

для каждого допустимого фиксированного δ. Исследуем основные свойства функций (1).

Лемма 1. *Многозначная функция* $\gamma(\alpha)$ *монотонно не возрастает, а* $\beta(\alpha)$ *монотонно не убывает в следующем смысле: если* α_1, α_2 — *такие произвольные положительные значения* α, *что* $\alpha_2>\alpha_1$, *то для любых соответствующих этим* $\alpha_{1,2}$ *экстремалей* $z^{\alpha_1}\in Z^{\alpha_1}$, $z^{\alpha_2}\in Z^{\alpha_2}$ *выполнены неравенства*

$$\Omega(z^{\alpha_1})\geqslant\Omega(z^{\alpha_2}),\quad I_\delta(z^{\alpha_1})\leqslant I_\delta(z^{\alpha_2}).$$

Доказательство. Возьмем произвольные α_1, α_2 $(0<\alpha_1<\alpha_2)$ и любые $z^{\alpha_1}\in Z^{\alpha_1}$, $z^{\alpha_2}\in Z^{\alpha_2}$. Поскольку z^{α_1} — решение задачи (4.2), то

$$\varphi(\alpha_1)=M^{\alpha_1}[z^{\alpha_1}]=\alpha_1\gamma(\alpha_1)+\beta(\alpha_1)\leqslant M^{\alpha_1}[z^{\alpha_2}]=$$
$$=\alpha_1\gamma(\alpha_2)+\beta(\alpha_2)=\varphi(\alpha_2)+(\alpha_1-\alpha_2)\gamma(\alpha_2). \quad (2)$$

Отсюда $(\alpha_2-\alpha_1)\gamma(\alpha_2)\leqslant\varphi(\alpha_2)-\varphi(\alpha_1)$. Аналогично можно получить неравенство $(\alpha_1-\alpha_2)\gamma(\alpha_1)\leqslant\varphi(\alpha_1)-\varphi(\alpha_2)$. Из этих двух неравенств имеем

$$(\alpha_2-\alpha_1)\Omega(z^{\alpha_2})=(\alpha_2-\alpha_1)\gamma(\alpha_2)\leqslant\varphi(\alpha_2)-\varphi(\alpha_1)\leqslant$$
$$\leqslant(\alpha_2-\alpha_1)\gamma(\alpha_1)=(\alpha_2-\alpha_1)\Omega(z^{\alpha_1}),\ \ \alpha_2>\alpha_1, \quad (3)$$

т. е. $\Omega(z^{\alpha_2})\leqslant\Omega(z^{\alpha_1})$ для любых $z^{\alpha_2}\in Z^{\alpha_2}$, $z^{\alpha_1}\in Z^{\alpha_1}$. Это доказывает монотонность функции $\gamma(\alpha)$. Тогда монотонность функции $\beta(\alpha)$ получается из неравенства

$$\beta(\alpha_2)-\beta(\alpha_1)=I_\delta(z^{\alpha_2})-I_\delta(z^{\alpha_1})\geqslant\alpha_1[\gamma(\alpha_1)-\gamma(\alpha_2)]=\alpha_1[\Omega(z^{\alpha_1})-\Omega(z^{\alpha_2})]\geqslant 0,$$

представляющего собой следствие из (2) и доказанной монотонности функции $\gamma(\alpha)$.

Лемма 2. *Функция* $\varphi(\alpha)$ *однозначна и непрерывна при* $\alpha>0$.

Доказательство. Для того чтобы убедиться в том, что $\varphi(\alpha)$ однозначна, достаточно заметить, что в силу (4.2)

$$\varphi(\alpha)=M^\alpha[z^\alpha]=\inf\{M^\alpha[z]\colon z\in D\}\quad\forall z^\alpha\in Z^\alpha.$$

Установим непрерывность функции $\varphi(\alpha)$ в произвольной точке $\alpha_0>0$. Для этого возьмем любые последовательности $\{\alpha_n^+\}$, $\{\alpha_n^-\}$, удовлетворяющие требованиям $\alpha_n^+\to\alpha_0+0$, $\alpha_n^-\to\alpha_0-0$ при $n\to\infty$. Тогда из (3) с учетом неравенства $\gamma(\alpha)=\Omega(z^\alpha)\geqslant\Omega^*$ получим

$$\varphi(\alpha_0)+(\alpha_n^+-\alpha_0)\Omega^*\leqslant\varphi(\alpha_0)+(\alpha_n^+-\alpha_0)\gamma(\alpha_n^+)\leqslant$$
$$\leqslant\varphi(\alpha_n^+)\leqslant\varphi(\alpha_0)+(\alpha_n^+-\alpha_0)\gamma(\alpha_0),$$
$$\varphi(\alpha_0)+(\alpha_0-\alpha_n^-)\Omega^*\leqslant\varphi(\alpha_0)+(\alpha_0-\alpha_n^-)\gamma(\alpha_n^-)\leqslant$$
$$\leqslant\varphi(\alpha_n^-)\leqslant\varphi(\alpha_0)+(\alpha_0-\alpha_n^-)\gamma(\alpha_0).$$

Переходя в этих неравенствах к пределу при $n\to\infty$, получаем $\varphi(\alpha_n^\pm)\to\varphi(\alpha_0)$. Отсюда следует, что функции $\varphi(\alpha)$ непрерывны в точке α_0 одновременно слева и справа.

Рассмотрим вопрос о непрерывности функций $\beta(\alpha)$, $\gamma(\alpha)$.

Лемма 3. *Функции* $\gamma(\alpha)$, $\beta(\alpha)$ *однозначны и непрерывны всюду, за исключением, быть может, не более чем счетного множества точек.*

Доказательство. Проведем его подробно только для функции $\beta_\delta(\alpha)$. Выберем при каждом $\alpha>0$ из всех значений $\{\beta_\delta(\alpha)\}$ этой многозначной функции, отвечающих данному α, одно определенное значение $\tilde\beta_\delta(\alpha)$. Полученную так уже однозначную функцию $\tilde\beta_\delta(\alpha)$ назовем *сечением* функции $\beta_\delta(\alpha)$. Пусть $\tilde\beta_\delta(\alpha)$, $\bar\beta_\delta(\alpha)$ — два различных сечения функции $\beta_\delta(\alpha)$. Как следствие леммы 1 эти функции будут монотонно не убывающими при $\alpha>0$. Как известно [93], такие функции имеют не более чем счетное число точек разрыва. Покажем, что если α_0 — точка непрерывности функции $\bar\beta_\delta(\alpha)$, то α_0 будет также и точкой непрерывности другого сечения $\tilde\beta_\delta(\alpha)$, причем $\bar\beta_\delta(\alpha_0)=\tilde\beta_\delta(\alpha_0)$.

Действительно, пусть, например, $\bar\beta_\delta(\alpha_0)>\tilde\beta_\delta(\alpha_0)$. Выберем последовательность $\{\alpha_n\}$ так, чтобы $\alpha_n\uparrow\alpha_0$. Согласно лемме 1, имеем $\beta_\delta(\alpha_n)\leqslant\beta_\delta(\alpha_0)$ как для многозначной функции. Поэтому и для различных сечений $\bar\beta_\delta(\alpha)$, $\tilde\beta_\delta(\alpha)$ выполнено неравенство $\bar\beta_\delta(\alpha_n)\leqslant\tilde\beta_\delta(\alpha_0)$. Отсюда с учетом непрерывности $\bar\beta_\delta(\alpha)$ в точке α_0 получим $\bar\beta_\delta(\alpha_0)\leqslant\tilde\beta_\delta(\alpha_0)$, что противоречит предположению. Значит, $\bar\beta_\delta(\alpha_0)=\tilde\beta_\delta(\alpha_0)$.

Далее, если, например, $\{\alpha_k\}_{k=1}^{\infty}$ — любая монотонно сходящаяся сверху к α_0 последовательность, то с учетом леммы 1 имеет место неравенство $\bar\beta_\delta(\alpha_{k+1})\leqslant\tilde\beta_\delta(\alpha_k)\leqslant\bar\beta_\delta(\alpha_{k-1})$ для любых $k>1$. Используя это, а также непрерывность $\bar\beta_\delta(\alpha)$ при $\alpha=\alpha_0$, приходим к выводу, что

$$\lim_{\alpha\to\alpha_0+0}\tilde\beta_\delta(\alpha)=\lim_{\alpha\to\alpha_0+0}\bar\beta_\delta(\alpha)=\bar\beta_\delta(\alpha_0)=\tilde\beta_\delta(\alpha_0),$$

т. е. $\tilde\beta_\delta(\alpha)$ непрерывна справа в точке α_0. Аналогично доказывается непрерывность слева.

Таким образом, все сечения многозначной функции $\beta_\delta(\alpha)$ совпадают в их общих точках непрерывности, т. е. всюду, за исключением, быть может, не более чем счетного числа точек. Различаются эти сечения тем самым лишь в их общих точках разрыва первого рода. Эти точки разрыва одновременно являются точками многозначности функции $\beta_\delta(\alpha)$. Аналогичное доказательство проводится и для $\gamma_\delta(\alpha)$.

Из леммы 3 можно сделать следующие выводы.

Следствие 1. *Если функции* $\beta_\delta(\alpha)$ *или* $\gamma_\delta(\alpha)$ *имеют точку разрыва* (*многозначности*) $\alpha_0>0$, *то соответствующее множество* $Z_\delta^{\alpha_0}$ *решений задачи* (4.2) *содержит по крайней мере две экстремали сглаживающего функционала* $M_\delta^\alpha[z]$. *Если экстремаль задачи* (4.2) *для данного* α *единственна, то точка* α *есть точка непрерывности функций* $\beta(\alpha)$, $\gamma(\alpha)$.

Следствие 2. *Функции* $\beta(\alpha)$, $\gamma(\alpha)$ *имеют общие точки разрыва* (*многозначности*).

Изучим поведение вспомогательных функций $\beta(\alpha)$, $\gamma(\alpha)$, $\varphi(\alpha)$ при $\alpha\to+0$ и при $\alpha\to+\infty$.

Лемма 4. *Существуют пределы*

$$\lim_{\alpha\to+0}\varphi(\alpha)=\lim_{\alpha\to+0}\beta(\alpha)=I_\delta^*\equiv\inf\{f[J_\delta(z)]\colon z\in D\}=f(J_\delta^*),$$

$$\lim_{\alpha\to+\infty}\frac{\varphi(\alpha)}{\alpha}=\lim_{\alpha\to+\infty}\gamma(\alpha)=\Omega^*\equiv\inf\{\Omega(z)\colon z\in D\},\qquad \lim_{\alpha\to+\infty}\frac{\beta(\alpha)}{\alpha}=0.$$

Существует также предел $\lim\limits_{\alpha\to+0}\gamma(\alpha)\equiv\Omega_\delta$ $(\Omega_\delta\geqslant\Omega^*)$.

Доказательство. Наиболее просто доказывается существование последнего из перечисленных пределов (конечного или бесконечного). Оно следует из обоснованного в лемме 1 монотонного невозрастания функции $\gamma(\alpha)$. Исследуем теперь поведение функций $\varphi(\alpha)$, $\beta(\alpha)$ при

$\alpha \to +0$. Для этого зададим произвольное $\varepsilon>0$. Тогда по определению точной нижней грани $I_\delta^* \equiv f[J_\delta^*]=\inf\{f[J_\delta(z)]: z\in D\}$ найдется такой элемент $z_\varepsilon \in D$, для которого $I_\delta^* \leqslant I_\delta(z_\varepsilon) \leqslant I_\delta^*+\varepsilon$. Поэтому можно написать неравенство

$$I_\delta^*+\alpha\Omega^* \leqslant I_\delta(z^\alpha)+\alpha\Omega(z^\alpha)=\varphi(\alpha)=M^\alpha[z^\alpha]\leqslant M^\alpha[z_\varepsilon]=$$
$$=\alpha\Omega(z_\varepsilon)+I_\delta(z_\varepsilon)\leqslant\alpha\Omega(z_\varepsilon)+I_\delta^*+\varepsilon,$$

в котором использован тот факт, что z^α есть экстремаль задачи (4.2). Если теперь перейти в этой цепочке неравенств к пределу при $\alpha \to +0$, получим

$$I_\delta^* \leqslant \varliminf_{\alpha\to+0} \varphi(\alpha)=\varlimsup_{\alpha\to+0}\varphi(\alpha)\leqslant I_\delta^*+\varepsilon,$$

откуда в силу произвольности ε и следует, что $\varphi(\alpha)\to I_\delta^*$ при $\alpha\to+0$. Учитывая это, а также неравенство

$$I_\delta^*\leqslant I_\delta(z_\delta^\alpha)=\beta(\alpha)=\varphi(\alpha)-\alpha\gamma(\alpha)\leqslant\varphi(\alpha)-\alpha\Omega^*,$$

в котором мы осуществим переход к пределу при $\alpha\to+0$, получаем $\beta(\alpha)\to I_\delta^*$.

Обратимся теперь к исследованию вспомогательных функций при $\alpha\to+\infty$. Снова зададим произвольное $\varepsilon>0$. Тогда по определению точной нижней грани Ω^* найдется элемент $\tilde z_\varepsilon\in D$, для которого выполнено неравенство $\Omega^*\leqslant\Omega(\tilde z_\varepsilon)\leqslant\Omega^*+\varepsilon$. Используем его, а также экстремальность элемента z^α для получения следующей оценки:

$$\Omega^*\leqslant\gamma(\alpha)\equiv[\varphi(\alpha)-\beta(\alpha)]/\alpha\leqslant(M^\alpha[z^\alpha]-I_\delta^*)/\alpha\leqslant(M^\alpha[\tilde z_\varepsilon]-I_\delta^*)/\alpha=$$
$$=\Omega(\tilde z_\varepsilon)+[I_\delta(\tilde z_\varepsilon)-I_\delta^*]/\alpha\leqslant\Omega^*+\varepsilon+[I_\delta(\tilde z_\varepsilon)-I_\delta^*]/\alpha.$$

Здесь использовано также неравенство $\beta(\alpha)\geqslant I_\delta^*$, вытекающее из монотонного неубывания функции $\beta(\alpha)$ (см. лемму 1) и из только что доказанного предельного соотношения $\beta(+0)=I_\delta^*$. Переходя к пределу при $\alpha\to+\infty$, получаем, что

$$\Omega^*\leqslant\lim_{\alpha\to+\infty}\gamma(\alpha)\leqslant\Omega^*+\varepsilon,\quad \Omega^*\leqslant\lim_{\alpha\to+\infty}\varphi(\alpha)/\alpha\leqslant\Omega^*+\varepsilon,$$

$$\Omega\leqslant\lim_{\alpha\to+\infty}\varphi(\alpha)/\alpha-\lim_{\alpha\to+\infty}\beta(\alpha)/\alpha\leqslant\Omega^*+\varepsilon\quad\forall\varepsilon>0,$$

откуда ясно, что $\gamma(\alpha)\to\Omega^*$, $\varphi(\alpha)/\alpha\to\Omega^*$ и, следовательно, $\beta(\alpha)/\alpha\to0$ при $\alpha\to+\infty$. Лемма доказана.

Необходимо подчеркнуть, что при доказательстве лемм 1—4 использовался лишь факт разрешимости задачи (4.2) для каждого $\alpha>0$ и совсем не применялись свойства функционалов J_δ, Ω, сформулированные в § 2. Эти свойства приходится использовать для проведения более детального исследования вспомогательных функций.

Лемма 5. *Предположим, что выполнены условия на функционалы J_δ, Ω из § 2. Тогда, если $\alpha_0>0$ — точка разрыва функций $\gamma(\alpha)$, $\beta(\alpha)$, то найдутся такие экстремали $z_+^{\alpha_0}$, $z_-^{\alpha_0}\in Z^{\alpha_0}$ задачи* (4.2) *с* $\alpha=\alpha_0$, *что*

$$\gamma(\alpha_0+0)=\Omega(z_+^{\alpha_0})\leqslant\gamma(\alpha_0)\leqslant\Omega(z_-^{\alpha_0})=\gamma(\alpha_0-0),$$
$$\beta(\alpha_0-0)=I_\delta(z_-^{\alpha_0})\leqslant\beta(\alpha_0)\leqslant I_\delta(z_+^{\alpha_0})=\beta(\alpha_0+0).$$

Доказательство. Рассмотрим сначала функцию $\gamma(\alpha)$. В силу ее монотонности при $\alpha>0$ существует предел

$$\lim_{\alpha\to\alpha_0-0}\gamma(\alpha)=\gamma(\alpha_0-0)\equiv\gamma^*=\sup\{\Omega(z^{\alpha_0}):\ z^{\alpha_0}\in Z^{\alpha_0}\}. \tag{4}$$

Пусть теперь α таково, что $0<\alpha_1\leqslant\alpha\leqslant\alpha_0$, где α_1 фиксировано. Тогда по лемме 1 можно написать $\gamma(\alpha_1)\geqslant\Omega(z^\alpha)\geqslant\gamma^*$. Первое из этих неравенств означает, что элементы $\{z^\alpha\}$ $(\alpha\in[\alpha_1,\alpha_0])$ принадлежат множеству $\Omega_1\equiv\{z\in D:\ \Omega(z)\leqslant\gamma(\alpha_1)\}$, которое будет τ — секвенциально компактным (см. предположение 2 из § 2). Поэтому существует последовательность $\{\alpha_n\}\in[\alpha_1,\alpha_0]$, которая, монотонно возрастая, сходится к α_0 и для которой $z^{\alpha_n}\to z^*\in D$ при $n\to\infty$. При этом, исходя из условия τ-секвенциальной полунепрерывности снизу функционала Ω и из (4), можно получить соотношение

$$\gamma^*=\lim_{\alpha\to\alpha_0-0}\gamma(\alpha)=\lim_{n\to\infty}\gamma(\alpha_n)=\lim_{n\to\infty}\Omega(z^{\alpha_n})=\varliminf_{n\to\infty}\Omega(z^{\alpha_n})\geqslant\Omega(z^*).$$

Применим это соотношение в следующей оценке:

$$\varphi(\alpha_0)=\lim_{n\to\infty}\varphi(\alpha_n)=\lim_{n\to\infty}M^{\alpha_n}[z^{\alpha_n}]=\lim_{n\to\infty}\{\alpha_n\Omega(z^{\alpha_n})+I_\delta(z^{\alpha_n})\}\geqslant$$

$$\geqslant\alpha_0\gamma^*+I_\delta(z^*)\geqslant\alpha_0\Omega(z^*)+I_\delta(z^*)=M^{\alpha_0}[z^*], \tag{5}$$

для получения которой использованы также непрерывность функции $\varphi(\alpha)$ (см. лемму 2), а также τ-секвенциальная полунепрерывность снизу функционала I_δ. Помимо этой оценки из экстремальности элемента z^{α_0} получается неравенство $M^{\alpha_0}[z^*]\geqslant M^{\alpha_0}[z^{\alpha_0}]=\varphi(\alpha_0)$, так что $\varphi(\alpha_0)=M^{\alpha_0}[z^*]$.

Последнее равенство приводит к двум следствиям. Во-первых, из (5) имеем $\gamma^*=\Omega(z^*)$. Во-вторых, z^* представляет собой решение задачи (4.2) при $\alpha=\alpha_0$, т. е. $z^*\in Z^{\alpha_0}$. Тогда (4) переписывается в виде

$$\gamma(\alpha_0-0)=\gamma^*=\Omega(z^*)\geqslant\Omega(z^{\alpha_0})=\gamma(\alpha_0)\quad\forall z^{\alpha_0}\in Z^{\alpha_0},$$

так что элемент z^* можно принять в качестве искомого $z^{\alpha_0}_-$.

Существование экстремали $z^{\alpha_0}_+\in Z^{\alpha_0}$ доказывается аналогичным образом при рассмотрении правой полуокрестности $(\alpha_0,\alpha_1]$ точки α_0, из которой выбирается монотонно убывающая и стремящаяся к α_0 последовательность $\{\alpha_n\}$.

Неравенства для $\beta(\alpha)$, доказываемые в этой лемме, следуют из неравенств для $\gamma(\alpha)$ и из непрерывности и однозначности функции $\varphi(\alpha)$. Например,

$$I_\delta(z^{\alpha_0})=\varphi(\alpha_0)-\alpha_0\Omega(z^{\alpha_0})\leqslant\varphi(\alpha_0)-\alpha_0\Omega(z^{\alpha_0}_+)=$$

$$=\varphi(\alpha_0+0)-\alpha_0\gamma(\alpha_0+0)=\beta(\alpha_0+0)=M^{\alpha_0}[z^{\alpha_0}_+]-\alpha_0\Omega(z^{\alpha_0}_+)=$$

$$=I_\delta(z^{\alpha_0}_+).$$

Лемма доказана.

Отметим, что неравенства из леммы 5 переходят в равенства, если α_0 — точка непрерывности функций (1). Таким образом, из следствия 1 и леммы 5 вытекает

Следствие 3. *Если при каждом* $\alpha>0$ *задача* (4.2) *минимизации сглаживающего функционала на множестве* D *имеет единственное решение, то при выполнении предположений из* § 2 *функции* $\gamma(\alpha)$, $\beta(\alpha)$ *непрерывны.*

Лемма 6. *При выполнении предположений из* § 2 *оказывается, что* $\beta(\alpha)\to\nu_\delta$ *при* $\alpha\to+\infty$, *а функция* $\varphi(\alpha)$ *имеет асимптоту* $\zeta(\alpha)=\alpha\Omega^*+\nu_\delta$, *причем* $\varphi(\alpha)\leqslant\zeta(\alpha)$.

Доказательство. Существование предела функции $\beta(\alpha)$ при $\alpha\to+\infty$ следует из ее монотонности. Вычислим этот предел. Пусть $\alpha_n\to+\infty$ при $n\to\infty$. Тогда по лемме 4 имеем $\gamma(\alpha_n)\to\Omega^*$. Отсюда n следует, что последовательность $\{z^{\alpha_n}\}$ τ-секвенциально компактна (см. предположение 2 из § 2), так что найдется последовательность $z_k\equiv z^{\alpha_{n_k}}$ $(k=1,\ 2,\ \ldots)$, τ-сходящаяся к некоторому элементу $z^*\in D$ при $k\to\infty$. Используя τ-секвенциальную полунепрерывность снизу функционалов Ω, I_δ (см. § 2 и лемму 4.2), можно получить

$$\begin{gathered}\lim_{k\to\infty}\beta(\alpha_{n_k})=\lim_{k\to\infty}I_\delta(z_k)\geqslant I_\delta(z^*),\\ \Omega^*=\lim_{k\to\infty}\gamma(\alpha_{n_k})=\lim_{k\to\infty}\Omega(z_k)\geqslant\Omega(z^*).\end{gathered}\tag{6}$$

Согласно формуле (5.33), из последнего неравенства можно сделать вывод о том, что $\Omega(z^*)=\Omega^*$, т. е. $z^*\in Z_0$. Далее, по лемме 5.12 вспомогательная задача (5.34) разрешима. Если $\hat z_\delta\in Z_0$ — какое-либо ее решение, то

$$\begin{aligned}\nu_\delta&=\inf\{I_\delta(z)\colon z\in Z_0\}\leqslant I_\delta(z^*)\leqslant\lim_{k\to\infty}\beta(\alpha_{n_k})=\\ &=\lim_{k\to\infty}\{\varphi(\alpha_{n_k})-\alpha_{n_k}\gamma(\alpha_{n_k})\}\leqslant\lim_{k\to\infty}\{M^{\alpha_{n_k}}[z^{\alpha_{n_k}}]-\alpha_{n_k}\Omega(z^*)\}\leqslant\\ &\leqslant\lim_{k\to\infty}\{M^{\alpha_{n_k}}[\hat z_\delta]-\alpha_{n_k}\Omega^*\}=\lim_{k\to\infty}\{\alpha_{n_k}\Omega(\hat z_\delta)-\alpha_{n_k}\Omega^*+I_\delta(\hat z_\delta)\}=I_\delta(\hat z_\delta)=\nu_\delta.\end{aligned}$$

В этой выкладке использована экстремальность элемента $z^{\alpha_{n_k}}$ как решения задачи (4.2), неравенства (6), а также следующие из определения решения задачи (5.34) равенства $\Omega(\hat z_\delta)=\Omega^*$, $I_\delta(\hat z_\delta)=\nu_\delta$. Таким образом, из этих неравенств можно сделать вывод, что

$$\lim_{k\to\infty}\beta(\alpha_{n_k})=\lim_{\alpha\to+\infty}\beta(\alpha)=\nu_\delta.$$

Для доказательства существования асимптоты $\zeta(\alpha)$ функции $\varphi(\alpha)$ при $\alpha\to+\infty$ заметим, что по лемме 4 имеем $\varphi(\alpha)/\alpha\to\Omega^*$ при $\alpha\to+\infty$. Тогда из соотношения

$$\begin{aligned}\beta(\alpha)=\varphi(\alpha)-\alpha\gamma(\alpha)\leqslant\varphi(\alpha)-\alpha\Omega^*=M^\alpha[z^\alpha]-\alpha\Omega^*\leqslant M^\alpha[\hat z_\delta]-\alpha\Omega^*=\\ =\alpha[\Omega(\hat z_\delta)-\Omega^*]+I_\delta(\hat z_\delta)=I_\delta(\hat z_\delta)=\nu_\delta,\end{aligned}$$

получаемого по аналогии с приведенной выше выкладкой, в результате предельного перехода при $\alpha\to+\infty$ и с учетом доказанной сходимости $\beta(\alpha)\to\nu_\delta$ следует, что $\varphi(\alpha)-\alpha\Omega^*\to\nu_\delta$. Из этого же соотношения получается доказываемое неравенство $\varphi(\alpha)\leqslant\alpha\Omega^*+\nu_\delta=\zeta(\alpha)$.

При построении регуляризующих алгоритмов в этой книге будут широко использоваться некоторые специальные решения уравнений, содержащих монотонные функции. В связи с этим приведем некоторые определения и вспомогательные результаты.

Предположим, что при $\alpha>0$ задана монотонно не убывающая (или не возрастающая) функция $\chi(\alpha)$. Как известно, такая функция будет непрерывна при $\alpha>0$ всюду, кроме не более чем счетного числа точек разрыва первого рода, если они имеются. Будем считать, что в этих точках разрыва функция $\chi(\alpha)$ является неоднозначной: если $\alpha^*>0$ — одна из точек разрыва, то множество значений $\chi(\alpha^*)$ состоит по меньшей мере из двух чисел $\chi(\alpha^*\pm 0)$. Типичными примерами такой функции $\chi(\alpha)$ служат вспомогательные функции (1).

Определение 1. Будем говорить, что число $\alpha_0\geqslant 0$ есть *решение уравнения*

$$\chi(\alpha)=0 \tag{7}$$

с монотонно не убывающей (не возрастающей) функцией $\chi(\alpha)$, если для любого $\alpha>0$ выполнено неравенство

$$(\alpha-\alpha_0)\chi(\alpha)\geqslant 0 \quad (\leqslant 0).$$

Определение 2. Если решение $\alpha_0>0$ уравнения (7) удовлетворяет равенству $\chi(\alpha_0)=0$, то назовем такое α_0 *обычным решением* рассматриваемого уравнения. Если же решение не является обычным, то будем называть его обобщенным.

Изучим некоторые свойства решений и обобщенных решений уравнения (7).

Лемма 7. 1) *Если монотонно не убывающая* (*не возрастающая*) *функция* $\chi(\alpha)$ *обладает свойством* $\chi_\infty\equiv\chi(+\infty)>0$ ($\chi_\infty<0$), *то уравнение* (7) *имеет по крайней мере одно решение* $\alpha_0\geqslant 0$. 2) *Если, кроме того,* $\chi_0\equiv\chi(+0)<0$ ($\chi_0>0$), *то любое решение уравнения* (7) *положительно.* 3) *Если* $\chi_\infty<0$ ($\chi_\infty>0$), *то уравнение* (7) *не имеет решений.*

Доказательство. Проведем его для случая монотонно не убывающей $\chi(\alpha)$. Из условия $\chi(+\infty)>0$ ясно, что при достаточно больших α выполнено неравенство $\chi(\alpha)>0$, так что множество $A_0=\{\alpha\colon \alpha>0,\ \chi(\alpha)\geqslant 0\}$ непусто. Поэтому существует число $\alpha_0\equiv\inf\{\alpha\colon \alpha\in A_0\}$. Возможны два случая.

В первом из них число α_0 равно нулю. Тогда для всех $\alpha>0$ функция $\chi(\alpha)$ будет неотрицательна, так что $\alpha_0=0$ удовлетворяет определению решения уравнения (7): $(\alpha-\alpha_0)\chi(\alpha)=\alpha\chi(\alpha)\geqslant 0$ для всякого $\alpha>0$.

Во втором случае $\alpha_0>0$. Тогда для $\alpha>\alpha_0$ выполнено неравенство $\chi(\alpha)\geqslant 0$, а для $\alpha\in(0,\alpha_0)$ — противоположное неравенство. При этом снова $(\alpha-\alpha_0)\chi(\alpha)\geqslant 0$ для каждого $\alpha>0$, и, следовательно, α_0 снова представляет собой решение уравнения (7). Заметим, что при выполнении условия $\chi_0<0$ функция $\chi(\alpha)$ отрицательна в некоторой окрестности нуля, так что может реализовываться лишь второй из рассмотренных случаев: $\alpha_0>0$, и это доказывает справедливость второй части леммы.

Рассматривая третью часть леммы, заметим, что из условия $\chi_\infty<0$ и из монотонного не убывания функции $\chi(\alpha)$ ясно, что $\chi(\alpha)<0$

для всякого $\alpha>0$. Поэтому, каково бы ни было $\alpha_0\geqslant 0$, неравенство $(\alpha-\alpha_0)\chi(\alpha)\geqslant 0$ не может выполняться для всех $\alpha>0$. Лемма доказана.

Следствие 4. *Если $\chi(\alpha)$ монотонно не убывает (не возрастает), то уравнение (7) имеет решение $\alpha_0=0$ в том и только в том случае, если $\chi_0\geqslant 0$ $(\chi_0\leqslant 0)$.*

Лемма 8. *Если $\alpha_0>0$ — обобщенное решение уравнения (7) с монотонно не убывающей (не возрастающей) функцией $\chi(\alpha)$, то α_0 представляет собой точку разрыва первого рода этой функции, причем $\chi(\alpha_0-0)<0$, $\chi(\alpha_0+0)>0$ $(\chi(\alpha_0-0)>0$, $\chi(\alpha_0+0)<0)$.*

Доказательство. Заметим, что из определения решения уравнения (7) с монотонно не убывающей функцией $\chi(\alpha)$ вытекают неравенства $\chi(\alpha)\leqslant 0$ при $0<\alpha<\alpha_0$ и $\chi(\alpha)\geqslant 0$ при $\alpha>\alpha_0$. Путем предельного перехода в этих неравенствах легко получить

$$\lim_{\alpha\to\alpha_0-0}\chi(\alpha)=\chi(\alpha_0-0)\leqslant 0,\qquad \lim_{\alpha\to\alpha_0+0}\chi(\alpha)=\chi(\alpha_0+0)\geqslant 0,\tag{8}$$

причем существование этих пределов следует из монотонности функции χ. Если предположить, что α_0 является точкой непрерывности функции $\chi(\alpha)$, то из (8) следует неравенство $0\leqslant\chi(\alpha_0+0)=\chi(\alpha_0)=\chi(\alpha_0-0)\leqslant 0$, т. е. α_0 будет представлять собой обычное решение уравнения (7). Однако это противоречит условию леммы 8. Поэтому α_0 будет точкой разрыва первого рода монотонной функции $\chi(\alpha)$. Согласно предположению о свойствах функции $\chi(\alpha)$ в точках разрыва, числа $\chi(\alpha_0\pm 0)$ входят в множество значений функции при $\alpha=\alpha_0$. Поэтому неравенства (8) не могут обращаться в равенства. Действительно, если, например, $\chi(\alpha_0-0)=0$, то α_0 снова будет обычным решением, что противоречит условию леммы. Тем самым выполнены неравенства $\chi(\alpha_0-0)<0$, $\chi(\alpha_0+0)>0$. Лемма доказана.

Из леммы 8, а также из определения 2 получаем

Следствие 5. *Если $\alpha_0>0$ — решение уравнения (7) с монотонно не убывающей (не возрастающей) функцией $\chi(\alpha)$, то $\chi(\alpha_0-0)\leqslant 0$, $\chi(\alpha_0+0)\geqslant 0$ $(\chi(\alpha_0-0)\geqslant 0$, $\chi(\alpha_0+0)\leqslant 0)$.*

Следствие 6. *Если $\alpha_0>0$ — обобщенное решение уравнения (7), то оно единственно.*

Чтобы убедиться в этом, предположим для определенности, что $\chi(\alpha)$ монотонно не убывает и уравнение (7) имеет два различных обобщенных решения α_0, α_1 $(0<\alpha_0<\alpha_1)$. Тогда в силу леммы 8 и монотонности функции $\chi(\alpha)$ выполнены противоречивые неравенства $0<\chi(\alpha_0+0)\leqslant\chi(\alpha_1-0)<0$, что и доказывает следствие 6. Понятно, что когда уравнение (7) имеет обычное решение, то оно в принципе может быть не единственным. В этом легко убедиться, рассматривая даже простейшие примеры.

Будем считать, что в случае, когда χ_0 — конечное число, функция $\chi(\alpha)$ доопределена по непрерывности в точке $\alpha=0$: $\chi(0)=\chi_0$.

Лемма 9. *Пусть уравнение (7) с монотонной функцией $\chi(\alpha)$, удовлетворяющей условиям $\chi(\alpha)\not\equiv 0$, $\chi_\infty\neq 0$, имеет не единственное решение. Тогда все эти решения являются точками некоторого отрезка $[\alpha_{\min}, \alpha_{\max}]$, где $\alpha_{\min}\geqslant 0$. Других решений уравнение (7) не имеет.*

Доказательство. Из следствия 6 ясно, что все решения уравнения (7) являются обычными, т. е. удовлетворяют равенству $\chi(\alpha)=0$. Поэтому из монотонности функции $\chi(\alpha)$ следует, что наряду с каждой парой решений α_1, $\alpha_2>0$ уравнение (7) имеет в качестве решений все точки отрезка $[\alpha_1, \alpha_2]$. Введем теперь числа $\alpha_{\min}=\inf\{\alpha>0\colon \chi(\alpha)=0\}$, $\alpha_{\max}=\sup\{\alpha>0\colon \chi(\alpha)=0\}$. Эти числа также будут обычными решениями уравнения (7), так как функция $\chi(\alpha)$ включает в множество своих значений величины $\chi(\alpha_0\pm 0)$ для всякого $\alpha_0>0$, а

$$\chi(\alpha_{\min}+0)=\lim_{\alpha\to\alpha_{\min}+0}\chi(\alpha)=0,\quad \chi(\alpha_{\max}-0)=\lim_{\alpha\to\alpha_{\max}-0}\chi(\alpha)=0.$$

Число $\alpha_{\max}$ конечно в силу требования $\chi_\infty\neq 0$. При этом, как отмечено выше, весь отрезок $[\alpha_{\min}, \alpha_{\max}]$ состоит из обычных решений уравнения (7), а вне этого отрезка таких решений уравнение иметь не может.

Если в лемме 9 заменить условие $\chi_\infty\neq 0$ на равенство $\chi_\infty=0$, то может оказаться, что $\alpha_{\max}=+\infty$.

На основе понятия решения уравнения с монотонной функцией можно установить некоторые дополнительные свойства вспомогательной функции $\varphi(\alpha)$. Для этого рассмотрим уравнение

$$\gamma(\alpha)=0,\quad \alpha>0. \tag{8'}$$

Функция $\gamma(\alpha)$ монотонно не возрастает (лемма 1), и по лемме 4 $\gamma_\infty\equiv\gamma(+\infty)=\Omega^*$. Поэтому разрешимость уравнения (8′) по лемме 7 связана со знаком величины Ω^*. Так, при $\Omega^*<0$ это уравнение разрешимо, и мы будем считать, что множество его решений есть отрезок $[\alpha_{\min}, \alpha_{\max}]$, где $\alpha_{\max}\geqslant\alpha_{\min}\geqslant 0$ (см. лемму 9). В случае $\Omega^*=0$ будем предполагать, что уравнение (8′) разрешимо и $\alpha_{\min}>0$ — его наименьшее решение.

Лемма 10. 1) *Если* $\Omega^*>0$, *то функция* $\varphi(\alpha)$ *монотонно возрастает и* $\varphi(+\infty)=+\infty$. 2) *Если* $\Omega^*<0$, *то* $\varphi(\alpha)$ *возрастает при* $0\leqslant\alpha\leqslant\alpha_{\min}$, *убывает при* $\alpha\geqslant\alpha_{\max}$, *а при* $\alpha_{\min}\leqslant\alpha\leqslant\alpha_{\max}$ *она постоянна.* 3) *Если* $\Omega^*=0$, *то функция* $\varphi(\alpha)$ *возрастает при* $0<\alpha\leqslant\alpha_{\min}$, *а при* $\alpha\geqslant\alpha_{\min}$ *она постоянна*: $\varphi(\alpha)=\beta(\alpha)\equiv\varphi_\infty=\nu_\delta$.

Доказательство. Оно основано на использовании неравенства (3). В первой части леммы для любых α_1, α_2 таких, что $\alpha_2>\alpha_1>0$, получим

$$\varphi(\alpha_2)-\varphi(\alpha_1)\geqslant\gamma(\alpha_2)(\alpha_2-\alpha_1)\geqslant\Omega^*(\alpha_2-\alpha_1)>0,$$

откуда и следуют утверждения леммы. При рассмотрении второй части необходимо заметить, что из невозрастания функции $\gamma(\alpha)$ и из леммы 9 следуют неравенства $\gamma(\alpha)>0$ при $0<\alpha<\alpha_{\min}$; $\gamma(\alpha)<0$ при $\alpha>\alpha_{\max}$. Поэтому из (3) следует

$$0<(\alpha_2-\alpha_1)\gamma(\alpha_2)\leqslant\varphi(\alpha_2)-\varphi(\alpha_1)\quad \forall\alpha_1,\alpha_2\colon 0<\alpha_1<\alpha_2<\alpha_{\min};$$

$$0>(\alpha_2-\alpha_1)\gamma(\alpha_1)\geqslant\varphi(\alpha_2)-\varphi(\alpha_1)\quad \forall\alpha_1,\alpha_2\colon \alpha_2>\alpha_1>\alpha_{\max}.$$

Если же α_1, α_2 — произвольные числа из отрезка $[\alpha_{\min}, \alpha_{\max}]$, то $\gamma(\alpha_1)=\gamma(\alpha_2)=0$, и из (3) получим $\varphi(\alpha_2)=\varphi(\alpha_1)$. Наконец, третья часть леммы получается так же, как и вторая, если формально положить $\alpha_{max}=+\infty$. В этом случае функция $\varphi(\alpha)$ постоянна при $\alpha\geqslant\alpha_{min}$, и, значит, по лемме 6 $\varphi(\alpha)=\nu_{\delta}$.

Проиллюстрируем свойства вспомогательных функций на простом примере.

Пример 1. Пусть $Z=\mathbf{R}$, а топология в Z определяется нормой $\|z\|=|z|$ для всех $z\in\mathbf{R}$. Зададим точный функционал $J(z)=|z-1|$ на

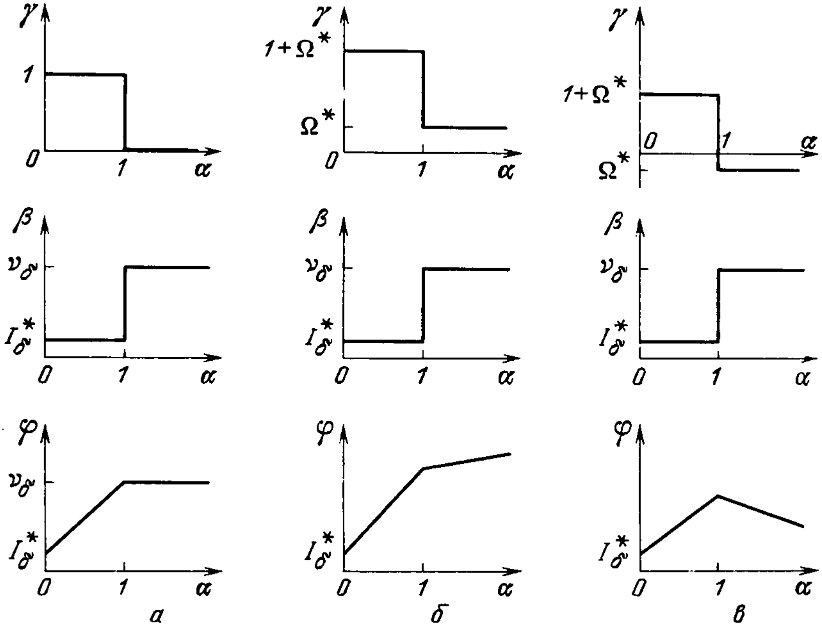

Рис. 3. Графики вспомогательных функций для: *а* — $\Omega^*=0$; *б* — $\Omega^*>0$; *в* — $\Omega^*<0$

множестве $D=\{z\in\mathbf{R}\colon\ z\geqslant 0\}$, а приближенный функционал определим так: $J_{\delta}(z)=J(z)+\delta$ $(\delta\geqslant 0)$. Будем считать, что $\Omega(z)=|z|+\Omega^*$, где число Ω^* может иметь различные значения. Наконец, предположим, что $f(x)=x$. Тогда сглаживающий функционал для такой задачи имеет вид $M^{\alpha}(z)=\alpha|z|+|z-1|+\delta+\alpha\Omega^*$, а экстремальная задача типа (4.2) о нахождении величин z^{α} записывается так:

$$z^{\alpha}=\operatorname{Arg\,inf}\{\alpha|z|+|z-1|+\alpha\Omega^*+\delta\colon\ z\geqslant 0\}. \tag{9}$$

Ее решения легко находятся: $z^{\alpha}=\{1$ при $0<\alpha<1$; $[0,1]$ при $\alpha=1$; 0 при $\alpha>1\}$. Отметим, что при $\alpha=1$ решения задачи (9) заполняют весь отрезок $[0, 1]$. Вычислим вспомогательные функции:

$$\gamma(\alpha)=\Omega(z^{\alpha})=\{1+\Omega^*\colon\ 0<\alpha<1;\ [\Omega^*,(1+\Omega^*)]\colon\ \alpha=1;\ \Omega^*\colon\ \alpha>1\},$$

$$\beta(\alpha)=\{\delta\colon\ 0<\alpha<1;\ [\delta,\delta+1]\colon\ \alpha=1;\ \delta+1\colon\ \alpha>1\},$$

$$\varphi(\alpha)=\{\alpha(1+\Omega^*)+\delta\colon\ 0<\alpha<1;\ 1+\Omega^*+\delta\colon\ \alpha=1;\ 1+\alpha\Omega^*+\delta\colon\ \alpha>1\}.$$

Их графики приведены на рис. 3 для различных Ω^*. В данном примере функции $\gamma(\alpha), \beta(\alpha)$ имеют точку разрыва $\alpha=1$, связанную с неоднозначностью решения задачи (9) при этом α. При этом значения $\{\gamma(1)\}$, $\{\beta(1)\}$ представляют собой целый отрезок.

В главе 2 будут приведены и другие примеры графиков вспомогательных функций.

Замечание. Однозначные функции, аналогичные функциям (1), впервые введены в работах [59, 61, 62, 149—152] для решения линейных операторных уравнений первого рода. В дальнейшем они изучались в ряде работ [80, 103—105, 142, 158—160, 206, 222—226]. Многозначные функции типа (1), возникающие при решении вариационным методом нелинейных операторных уравнений, а также экстремальных задач, изучались в [119—125, 130, 133, 134].

§ 7. Обобщенный принцип невязки для экстремальных задач

В данном параграфе будет изучаться алгоритм устойчивого приближенного нахождения Ω-оптимальных решений экстремальной задачи (2.2), который строится по схеме § 4 и использует для выбора параметра регуляризации и для отбора экстремали функции

$$\pi(\alpha)\equiv f\left[\Psi\left(\delta,\Omega(z_\delta^\alpha)\right)+\lambda_\delta\right]\equiv\Pi_\delta\left[z_\delta^\alpha\right], \tag{1}$$

$$\rho(\alpha)\equiv\beta(\alpha)-\pi(\alpha)\equiv P_\delta\left[z_\delta^\alpha\right]=I_\delta\left[z_\delta^\alpha\right]-\Pi_\delta\left[z_\delta^\alpha\right]. \tag{2}$$

В дальнейшем функцию $\rho(\alpha)$ будем называть *обобщенной невязкой.* В гл. 2 будет показано, что при решении операторных уравнений вариационным методом по схеме сведения к задаче (2.2) функция $\rho(\alpha)$ является естественным обобщением понятия невязки уравнения. Эта же терминология используется нами и при решении экстремальных задач.

Отметим, что для вычисления функций (1), (2) необходимо найти решение λ_δ задачи первого типа (3.2), а также найти экстремали z^α функционала (4.1) из задачи (4.2). На основании лемм 6.1—6.6 и с учетом свойств функций f, Ψ можно сделать следующие выводы о свойствах вспомогательных функций (1), (2) при фиксированном δ.

Теорема 1.1. *Функция* (1) *монотонно не возрастает, а функция* (2) *монотонно не убывает при* $\alpha>0$. *Эти функции непрерывны при* $\alpha>0$ *всюду, кроме, быть может, не более чем счетного множества их общих точек разрыва первого рода, являющихся одновременно их точками многозначности. Кроме того, существуют пределы*:

$$\lim_{\alpha\to+0}\rho(\alpha)=f(J_\delta^*)-f\left[\Psi(\delta,\Omega_\delta)+\lambda_\delta\right]\equiv\rho_0\equiv\rho_0^\delta,$$

$$\lim_{\alpha\to+\infty}\rho(\alpha)\equiv\rho_\infty=\rho_\infty^\delta,\quad \lim_{\alpha\to+0}\pi(\alpha)=f\left[\Psi(\delta,\Omega_\delta)+\lambda_\delta\right],$$

$$\lim_{\alpha\to+\infty}\pi(\alpha)=f\left[\Psi(\delta,\Omega^*)+\lambda_\delta\right].$$

2. *Пусть выполнены предположения* 1—3 *из* § 2 *для функционалов* J_δ, Ω. *Тогда, если* $\alpha=\alpha_0>0$ *является точкой разрыва функций* (1), (2), *множество решений* Z^{α_0} *экстремальной задачи* (4.2) *минимизации*

сглаживающего функционала при $\alpha=\alpha_0$ *содержит по крайней мере два элемента* $z_+^{\alpha_0}$, $z_-^{\alpha_0}$. *При этом*

$$\rho(\alpha_0 \pm 0) = P_\delta[z_\pm^{\alpha_0}], \quad \pi(\alpha_0 \pm 0) = \Pi_\delta[z_\pm^{\alpha_0}].$$

Кроме того, из этих условий следует, что $\rho_\infty = \nu_\delta - f[\Psi(\delta, \Omega^*) + \lambda_\delta]$.

На основании неравенства $\lambda_\delta \geqslant J_\delta^*$, доказанного в теореме 3.1, используя неотрицательность функции Ψ и монотонное возрастание функции f, можно утверждать, что

$$\rho_0^\delta = f(J_\delta^*) - f[\Psi(\delta, \Omega_\delta) + \lambda_\delta] \leqslant f(\lambda_\delta) - f[\Psi(\delta, \Omega_\delta) + \lambda_\delta] \leqslant 0.$$

Рассмотрим теперь уравнение

$$\rho(\alpha) = 0 \tag{3}$$

с монотонной функцией ρ. Из леммы 6.7 вытекает

Теорема 2. *Если для рассматриваемого фиксированного* δ *справедливо неравенство* $\rho_\infty > 0$, *то уравнение* (3) *имеет по крайней мере одно решение* $\alpha_\delta \geqslant 0$. *Если, кроме того,* $\rho_0 < 0$, *то любое решение уравнения* (3) *положительно.*

Приведем достаточные условия, гарантирующие выполнение неравенств $\rho_0 < 0$, $\rho_\infty > 0$ при фиксированном δ.

Теорема 3. *Пусть выполнены предположения* 1–3 *из* §2 *и условия*: а) $Z^* \cap Z_0 = \varnothing$; б) *функция* $\Psi(\delta, \Omega)$ *монотонно возрастает по* Ω *при каждом* $\delta \in \mathbf{R}_+^k$ $(\delta \neq \theta)$. *Тогда существует такая константа* $\Delta_0 > 0$, *что при любом фиксированном* δ $(\|\delta\| < \Delta_0)$ *выполнены неравенства* $\rho_0^\delta < 0$, $\rho_\infty^\delta > 0$.

Доказательство. Прежде всего докажем, что предположения 1—3 вместе с требованием $Z^* \cap Z_0 = \varnothing$ ведут к неравенству

$$\nu_0 \equiv \inf\{I(z)\colon z \in Z_0\} > \inf\{I(z)\colon z \in D\} = f(J^*) \equiv \mu_0, \tag{4}$$

где $I(z) \equiv f[J(z)]$. Из предположений 1, 2 из § 2 и леммы 5.12 ясно, что множество Z_0 будет τ-секвенциально компактным. Тогда функционал $I(z)$, являясь по лемме 4.1 τ-секвенциально полунепрерывным снизу на D, будет также обладать этим же свойством и на Z_0. Следовательно, по теореме типа Вейерштрасса (см. § 0.1) экстремальная задача

$$\operatorname{Arg\,inf}\{I(z)\colon z \in Z_0\}$$

разрешима, и существует такой элемент $z_0 \in Z_0$, для которого $I(z_0) = \nu_0$. Если предположить, что неравенство (4) неверно, то будет выполняться равенство $\nu_0 = \mu_0$. Но это будет означать, что

$$f(J^*) = \inf\{I(z)\colon z \in D\} = f[J(z_0)],$$

и, следовательно, z_0 окажется решением задачи (2.2): $z_0 \in Z^*$, и одновременно $z_0 \in Z_0$. Это противоречит требованию теоремы 3 о том, что $Z^* \cap Z_0 = \varnothing$. Таким образом, неравенство (4) выполнено.

Докажем теперь, что

$$\nu_\delta \equiv \inf\{I_\delta(z)\colon z \in Z_0\} \to \nu_0$$

при $\|\delta\|\to 0$. При выполнении предположений 1, 2 из § 2 задача (5.34) разрешима в силу леммы 5.12. Пусть $\hat{z}_\delta \in Z_0$ — какое-нибудь ее решение. Тогда, принимая во внимание условие аппроксимации (2.3) точного функционала J приближенным J_δ, возрастание функции f, а также равенства $\Omega(\hat{z}_\delta)=\Omega^*$, $\Omega(z_0)=\Omega^*$, вытекающие из включений $\hat{z}_\delta \in Z_0$, $z_0 \in Z_0$, получаем неравенства

$$I(z_0)=f[J(z_0)]=\nu_0=\inf\{I(z)\colon z\in Z_0\}\leqslant f[J(\hat{z}_\delta)]\leqslant f[J_\delta(\hat{z}_\delta)+\Psi(\delta,\Omega^*)],$$

$$I_\delta(\hat{z}_\delta)=f[J_\delta(\hat{z}_\delta)]=\nu_\delta=\inf\{I_\delta(z)\colon z\in Z_0\}\leqslant f[J_\delta(z_0)]\leqslant f[J(z_0)+\Psi(\delta,\Omega^*)].$$

Из этих соотношений, используя непрерывность и возрастание функции f, получаем

$$f[f^{-1}(\nu_0)-\Psi(\delta,\Omega^*)]\leqslant\nu_\delta\leqslant f[f^{-1}(\nu_0)+\Psi(\delta,\Omega^*)], \tag{5}$$

где f^{-1} — функция, обратная к f. Переходя в (5) к пределу при $\Delta=\|\delta\|\to 0$ и учитывая непрерывность функции f, а также свойства 3 а), в) функции Ψ (см. § 2), находим, что

$$\nu_0=f[f^{-1}(\nu_0)]\leqslant\lim_{\Delta\to 0}\nu_\delta\leqslant f[f^{-1}(\nu_0)]=\nu_0,$$

что и доказывает искомую сходимость.

Эта сходимость вместе с неравенством (4) позволяет обосновать, что $\rho_\infty^\delta>0$ для всех $\|\delta\|\leqslant\Delta_0$. Действительно, пусть это не так и для некоторой последовательности δ_n ($\|\delta_n\|\to 0$ при $n\to\infty$) оказывается, что $\rho_\infty^n\equiv\rho_\infty^{\delta_n}\leqslant 0$. Поскольку в п. 2 теоремы 1 утверждается, что $\rho_\infty^n=\nu_{\delta_n}-f[\Psi(\delta_n,\Omega^*)+\lambda_{\delta_n}]$, то, переходя к пределу в этом равенстве при $n\to\infty$, получаем

$$0\geqslant\lim_{n\to\infty}\rho_\infty^n=\lim_{n\to\infty}\nu_{\delta_n}-\lim_{n\to\infty}f[\Psi(\delta_n,\Omega^*)+\lambda_{\delta_n}]=\nu_0-f[J^*]=\nu_0-\mu_0>0.$$

Здесь использованы также сходимости $\Psi(\delta_n,\Omega^*)\to 0$, $\lambda_{\delta_n}\to J^*$ (см. теорему 3.1) при $\|\delta_n\|\to 0$. Это противоречивое неравенство доказывает, что $\rho_\infty^\delta>0$ при достаточно малых $\|\delta\|$.

Теперь докажем, что $\rho_0^\delta<0$ при достаточно малых $\|\delta\|$. Для этого заметим, что по теореме 1 и вследствие неравенства $J_\delta^*\leqslant\lambda_\delta$ (см. теорему 3.1) получается оценка

$$\rho_0^\delta=f[J_\delta^*]-f[\Psi(\delta,\Omega_\delta)+\lambda_\delta]=f[J_\delta^*]-f[\lambda_\delta]+$$
$$+\{f(\lambda_\delta)-f[\Psi(\delta,\Omega_\delta)+\lambda_\delta]\}\leqslant f[\lambda_\delta]-f[\Psi(\delta,\Omega_\delta)+\lambda_\delta].$$

Поэтому для доказательства искомого неравенства достаточно проверить, что $f(\lambda_\delta)<f[\Psi(\delta,\Omega_\delta)+\lambda_\delta]$ при $\|\delta\|\leqslant\Delta_0=\mathrm{const}$. Предположим, что это не так. Тогда найдется последовательность δ_n ($\|\delta_n\|\to 0$ при $n\to\infty$), для которой

$$f(\lambda_{\delta_n})=f[\Psi(\delta_n,\Omega_{\delta_n})+\lambda_{\delta_n}].$$

В силу возрастания функции f и неотрицательности Ψ отсюда следует, что $\Psi(\delta_n,\Omega_{\delta_n})=0$ для любого n. Из этого равенства получаются важные следствия.

Прежде всего оказывается, что $\Omega_{\delta_n}=\Omega^*$. Чтобы убедиться в этом, заметим, что из леммы 6.4 следует неравенство $\Omega_{\delta_n}\geqslant\Omega^*$. Если предположить выполнение строгого неравенства $\Omega_{\delta_n}>\Omega^*$, то вследствие возрастания функции Ψ по второму аргументу получим $0=\Psi(\delta_n, \Omega_{\delta_n})>\Psi(\delta_n, \Omega^*)\geqslant 0$. Это противоречивое неравенство доказывает, что $\Omega_{\delta_n}=\Omega^*$. Из невозрастания функции $\gamma(\alpha)$ и по лемме 6.4 тогда выполнено равенство $\Omega(z^{\alpha}_{\delta_n})=\gamma_{\delta_n}(\alpha)=\Omega_{\delta_n}=\Omega^*$, справедливое для любого $\alpha>0$. Используя это и неравенство (6.3), получаем

$$(\alpha_2-\alpha_1)\Omega^*=\varphi_{\delta_n}(\alpha_2)-\varphi_{\delta_n}(\alpha_1)=\alpha_2\Omega^*+\beta_{\delta_n}(\alpha_2)-\alpha_1\Omega^*-\beta_{\delta_n}(\alpha_1)\quad \forall\alpha_2>\alpha_1>0,$$

т. е. $\beta_{\delta_n}(\alpha_2)=\beta_{\delta_n}(\alpha_1)=\mathrm{const}$ для любого $\alpha_2>\alpha_1>0$. Тогда из лемм 6.4, 6.6 ясно, что

$$f\left[J^*_{\delta_n}\right]=\lim_{\alpha\to+0}\beta_{\delta_n}(\alpha)=\lim_{\alpha\to+\infty}\beta_{\delta_n}(\alpha)=\nu_{\delta_n}\quad \forall n.$$

Далее, из равенства $\Omega_{\delta_n}=\Omega^*$ вытекает, что $\Psi(\delta_n, \Omega_{\delta_n})=0=\Psi(\delta_n, \Omega^*)$, и это с учетом неравенства (5) дает $\nu_{\delta_n}=\nu_0$. Таким образом, получим $f(\lambda_{\delta_n})\geqslant f(J^*_{\delta_n})=\nu_{\delta_n}=\nu_0$ для любого n. Переходя в последнем неравенстве к пределу при $n\to\infty$ и принимая во внимание, что по теореме 3.1 $\lambda_{\delta_n}\to J^*$ при $n\to\infty$, получаем $f(J^*)\geqslant\nu_0$. Но это противоречит установленному выше неравенству (4). Полученное противоречие доказывает, что $\rho_0^\delta<0$ при $\|\delta\|\leqslant\Delta_0$.

Следствие 1. *При выполнении условий теоремы* 3 *и с учетом теоремы* 2 *уравнение* (3) *имеет при достаточно малых* $\|\delta\|$ *лишь положительные решения* α_δ.

Условие $Z^*\bigcap Z_0=\varnothing$ теоремы 3 означает, что среди решений исходной экстремальной задачи (2.2) нет элементов, реализующих глобальный минимум функционала $\Omega(z)$ на множестве D. Такое требование представляется вполне естественным, так как иначе задача о поиске Ω-оптимальных решений экстремальной задачи (2.2) становится более простой, сводясь для ряда задач к устойчивой минимизации функционала J_δ на множестве Z_0. Случай, когда $Z^*\bigcap Z_0\neq\varnothing$, более подробно исследован в § 10.

Другое условие теоремы 3, согласно которому мера аппроксимации $\Psi(\delta, \Omega)$ возрастает по второму аргументу при фиксированном первом $(\delta\neq\theta)$, выполнено обычно для широкого класса задач. Примеры будут приведены в § 11.

Согласно следствию 1, по лемме 6.9 уравнение (3) может иметь не единственное решение. Имеются, однако, простые условия, обеспечивающие единственность решения уравнения (3), даже если оно обычное.

Теорема 4. *Предположим, что выполнены предположения* 1—3 *из* § 2, $f\in\mathscr{F}^m$ *и, кроме того,* $\rho_0<0$, $\rho_\infty>0$ *для рассматриваемого* δ. *Пусть также задача минимизации сглаживающего функционала* (4.2) *имеет при каждом* $\alpha>0$ *единственное решение* z^α, *причем для произвольных различных* α_1, $\alpha_2>0$ *реализуется требование* $z^{\alpha_1}\neq z^{\alpha_2}$. *Тогда уравнение* (3) *имеет единственное обычное решение* $\alpha_\delta>0$.

Доказательство. Согласно следствию 6.1, в силу определения (1), (2) функций $\pi(\alpha)$, $\beta(\alpha)$ они будут непрерывны при $\alpha>0$. Тогда по теореме 2 и по определению 6.2 уравнение (3) имеет по крайней мере одно обычное решение $\alpha_1 \equiv \alpha_\delta > 0$. Покажем, что оно единственное. Предположим, что имеется еще решение $\alpha_2 > 0$ $(\alpha_2 \neq \alpha_1)$ уравнения (3). Тогда из следствия 6.6 ясно, что оно также обычное. Поскольку по условию теоремы экстремаль задачи (4.2) единственна при каждом α и $z^{\alpha_1} \neq z^{\alpha_2}$, то

$$\alpha_1 \Omega(z^{\alpha_1}) + I_\delta(z^{\alpha_1}) = M^{\alpha_1}[z^{\alpha_1}] < M^{\alpha_1}[z^{\alpha_2}] = \alpha_1 \Omega(z^{\alpha_2}) + I_\delta(z^{\alpha_2}). \qquad (6)$$

Учитывая, что $\alpha_{1,2}$ — обычные решения уравнения (3), т. е. что

$$\rho(\alpha_1) = \rho(\alpha_2) = I_\delta(z^{\alpha_{1,2}}) - \Pi_\delta(z^{\alpha_{1,2}}) = 0,$$

получаем из (6)

$$\Phi_1[\Omega(z^{\alpha_1})] \equiv \alpha_1 \Omega(z^{\alpha_1}) + \Pi_\delta(z^{\alpha_1}) < \alpha_1 \Omega(z^{\alpha_2}) + \Pi_\delta(z^{\alpha_2}) \equiv \Phi_1[\Omega(z^{\alpha_2})],$$

где $\Phi_1(\Omega) = \alpha_1 \Omega + f[\Psi(\delta, \Omega) + \lambda_\delta]$. Функция $\Phi_1(\Omega)$ монотонно возрастает, так как первое слагаемое, в котором $\alpha_1 > 0$, монотонно возрастает, а второе слагаемое, принимая во внимание свойства функций f, Ψ, монотонно не убывает. Поэтому последнее неравенство эквивалентно следующему: $\Omega(z^{\alpha_1}) < \Omega(z^{\alpha_2})$. Повторим теперь приведенные рассуждения, начиная с (6), поменяв местами в формулах величины α_1 и α_2. Тогда получим противоречащее установленному неравенство $\Omega(z^{\alpha_2}) < \Omega(z^{\alpha_1})$. Это и доказывает единственность решения.

Опишем теперь алгоритм приближенного нахождения Ω-оптимальных решений задачи (4.2), который называется *обобщенным принципом невязки* (о.п.н.) для экстремальных задач. Для этого необходимо указать метод выбора параметра регуляризации α и правило отбора экстремали (см. § 4).

Параметр регуляризации в алгоритме о.п.н. выбирается, как любое решение $\alpha_\delta > 0$ уравнения (3). Пусть $Z_\delta^{\alpha_\delta}$ — множество экстремалей задачи (4.2) минимизации сглаживающего функционала для $\alpha = \alpha_\delta$. Сформулируем правило отбора экстремали из этого множества. Для этого зададим константы $C > 1$, $q > 1$, обозначим $\alpha_1 \equiv \alpha_1(\delta) = \alpha_\delta / q$, $\alpha_2 \equiv \alpha_2(\delta) = \alpha_\delta q$ и возьмем две любые экстремали z^{α_1}, z^{α_2} задачи (4.2) для указанных α_1, α_2.

Правило отбора 1. Если выполнено неравенство

$$I_\delta[z^{\alpha_2(\delta)}] \geqslant C \Pi_\delta[z^{\alpha_1(\delta)}] - (C-1) f(\lambda_\delta), \qquad (7)$$

то из множества $Z_\delta^{\alpha_\delta}$ выбираем такую экстремаль z^{α_δ}, для которой справедливо неравенство

$$P_\delta[z^{\alpha_\delta}] \equiv I_\delta[z^{\alpha_\delta}] - \Pi_\delta[z^{\alpha_\delta}] \leqslant 0. \qquad (8)$$

Если же выполнено неравенство

$$I_\delta[z^{\alpha_2(\delta)}] \leqslant C \Pi_\delta[z^{\alpha_1(\delta)}] - (C-1) f(\lambda_\delta), \qquad (9)$$

то экстремаль z^{α_δ} выбираем так, чтобы выполнялось неравенство

$$P_\delta[z^{\alpha_\delta}] \geqslant 0. \qquad (10)$$

Полученный таким образом элемент $z^{\alpha_\delta} \in D$ примем в качестве приближения к множеству $\bar{Z}$ Ω-оптимальных решений задачи (2.2).

Теоремы 2—4 обосновывают возможность выбора параметра регуляризации в алгоритме о.п.н. В случае, если уравнение (3) имеет обычные решения, выполнено равенство

$$\rho(\alpha_\delta)=P_\delta\left[z^{\alpha_\delta}\right]=0 \quad \forall z^{\alpha_\delta}\in Z_\delta^{\alpha_\delta},$$

т. е. для любой экстремали z^{α_δ}, отвечающей выбранному параметру регуляризации, выполнены одновременно (9) и (10). Таким образом, в рассматриваемом случае нет необходимости использовать правило отбора. Иное положение складывается, если уравнение (3) имеет обобщенное решение, которое по следствию 6.6 будет единственным. Тогда в соответствии с леммой 6.8 и по теореме 1 точка $\alpha_\delta>0$ будет точкой разрыва (многозначности) монотонно не убывающей функции $\rho(\alpha)$, причем по лемме 6.5 найдутся такие экстремали $z_+^{\alpha_\delta}$, $z_-^{\alpha_\delta}$ задачи (4.2) с $\alpha=\alpha_\delta$, для которых

$$\rho(\alpha_\delta-0)=P_\delta\left[z_-^{\alpha_\delta}\right]<0, \quad \rho(\alpha_\delta+0)=P_\delta\left[z_+^{\alpha_\delta}\right]>0.$$

В этой ситуации использование правила отбора 1 становится необходимым, и ясно, что неравенство (8) реализуется, например, при $z^{\alpha_\delta}=z_-^{\alpha_\delta}$, а неравенство (10) — при $z^{\alpha_\delta}=z_+^{\alpha_\delta}$. Сказанное обосновывает тот факт, что правило отбора 1 гарантирует выбор элемента z^{α_δ}.

Докажем, что алгоритм о.п.н.— регуляризующий. Для этого предположим, что при каждом $\delta\in\mathbf{R}_+^k$ ($\|\delta\|\leqslant\Delta_0$) величины $\alpha_\delta>0$, z^{α_δ} находятся в соответствии с указанным принципом выбора параметра регуляризации и с правилом отбора 1. Будем считать также, что выполнены предположения 1—3 из § 2, а также свойства функции f из § 4.

Теорема 5. *Для произвольной последовательности* $\{\delta_n\}$ ($\delta_n\in\mathbf{R}_+^k$, $\delta_n\neq\theta$), *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательность приближений* $z_n=z^{\alpha_{\delta_n}}$, *полученных по алгоритму о.п.н.* τ-*сходится к* $\bar{Z}$ *при* $n\to\infty$, *причем* $\Omega(z_n)\to\bar{\Omega}$.

Доказательство. Легко видеть, что семейство $\{z^{\alpha_\delta}\}$ и вспомогательные семейства $\{z^{\alpha_{1,2}(\delta)}\}$, фигурирующие в правиле отбора алгоритма о.п.н., удовлетворяют требованиям а)—д), изложенным в § 5. Действительно, условие а) выполнено в силу неравенства

$$M_\delta^{\alpha_\delta}\left[z^{\alpha_\delta}\right]\leqslant M_\delta^{\alpha_\delta}[\bar{z}]=\alpha_\delta\Omega(\bar{z})+f\left[J_\delta(\bar{z})\right] \quad \forall\bar{z}\in\bar{Z}$$

и с учетом неравенства (5.2), если в а) положить $\bar{\Omega}_\delta\equiv\bar{\Omega}$. Аналогичным образом обосновывается и условие б), так как $\alpha_2(\delta)=\alpha_\delta q$. Условие в) следует из леммы 6.1, а условие г) — из правила отбора 1. Наконец, условие д) получается с учетом монотонности функции $\gamma(\alpha)$ следующим образом: если $\alpha_\delta\geqslant\alpha_0>0$, то $\alpha_1=\alpha_\delta/q\geqslant\alpha_0/q$ и

$$\Omega(z^{\alpha_1})=\gamma(\alpha_1)\leqslant\gamma(\alpha_0/q)=\Omega(z^{\alpha_0/q}).$$

Итак, выполнены условия лемм 5.7—5.9, следствия 5.7 и леммы 5.10, на основании которых можно сделать вывод о том, что

$$\lim_{\Delta\to0} I_\delta(z_\delta^{\alpha_\delta})=I^*, \tag{11}$$

$$\overline{\lim_{\Delta\to0}}\,\Omega(z_\delta^{\alpha_\delta})\leqslant\bar{\Omega}. \tag{12}$$

В свою очередь, соотношения (11), (12) представляют собой условия леммы 5.1 и ее следствие 5.1, из которых и получаются утверждения теоремы 5.

Тем самым показано, что алгоритм о.п.н. позволяет находить устойчивые приближения к множеству Ω-оптимальных решений экстремальной задачи (2.2).

Замечание. Обобщенный принцип невязки для экстремальных задач предложен в работах [124, 125, 130].

§ 8. Обобщенный принцип квазирешений для экстремальных задач

В § 7 исследован алгоритм о.п.н., позволяющий определять приближенное решение задачи (2.4) по данным $\{J_\delta, \Psi, f, \delta\}$. В некоторых случаях кроме этой информации имеются также данные о величине $\bar{\Omega}$. Как правило, это выражается в том, что известна оценка сверху для $\bar{\Omega}$. Поэтому представляется естественным использовать такую дополнительную информацию в структуре регуляризирующего алгоритма. Это можно сделать по схеме § 4, вводя оценку ω_δ величины $\bar{\Omega}$ в правило выбора параметра регуляризации и в правило отбора.

Рассмотрим уравнение с монотонно невозрастающей (см. лемму 6.1) функцией $\gamma(\alpha)$:

$$\gamma(\alpha)=\omega_\delta, \tag{1}$$

где ω_δ — заданное число, а δ фиксировано. Непосредственно из лемм 6.1, 6.4, 6.7 получается

Теорема 1. *Пусть число* ω_δ *удовлетворяет неравенству* $\Omega^*<\omega_\delta<\Omega_\delta$. *Тогда уравнение* (1) *имеет решение* $\alpha_\delta>0$.

В принципе такое решение может оказаться не единственным. Однако по аналогии с теоремой 7.4 можно установить, что верна

Теорема 2. *Предположим, что выполнены предположения* 1—3 *из* § 2, $f\in\mathscr{F}^m$ *и* $\Omega^*<\omega_\delta<\Omega_\delta$. *Пусть также задача* (4.2) *имеет единственное решение при каждом* $\alpha>0$, *причем* $z^{\alpha_1}\neq z^{\alpha_2}$ *для любых* $\alpha_1\neq\alpha_2$ $(\alpha_1, \alpha_2>0)$. *Тогда уравнение* (1) *имеет единственное обычное решение* $\alpha_\delta>0$.

Действительно, согласно следствию 6.1 из леммы 6.3, функция $\gamma(\alpha)$ будет непрерывна при $\alpha>0$, если выполнено условие единственности экстремали задачи (4.2). Далее по лемме 6.4 по условиям на число ω_δ можно сделать вывод, что для функции $\chi(\alpha)\equiv\gamma(\alpha)-\omega_\delta$ выполняются условия леммы 6.7. Тем самым доказана разрешимость уравнения $\chi(\alpha)=0$ в обычном смысле. Единственность его решения доказывается, как в теореме 7.4, исходя из неравенства (7.6). При этом снова предполагается, что имеется два обычных решения уравнения α_1, $\alpha_2>0$ $(\alpha_1\neq\alpha_2)$. Тогда из (7.6) и из равенств

$$\gamma(\alpha_1)=\gamma(\alpha_2)=\Omega(z^{\alpha_1})=\Omega(z^{\alpha_2})=\omega_\delta,$$

представляющих собой определения обычных решений α_1, α_2 уравнения (1), получаем

$$\alpha_1\omega_\delta+I_\delta(z^{\alpha_1})<\alpha_1\omega_\delta+I_\delta(z^{\alpha_2}),$$

т. е. $I_\delta(z^{\alpha_1})<I_\delta(z^{\alpha_2})$. Поменяв местами α_1 и α_2 в неравенстве (7.6), используя их равноправность, получаем таким же образом другое неравенство $I_\delta(z^{\alpha_2})<I_\delta(z^{\alpha_1})$. Противоречивость двух последних неравенств доказывает единственность обычного решения $\alpha_\delta>0$ уравнения (1).

Условие, налагаемое на число ω_δ в теоремах 1, 2, предполагает, что $\Omega^*<\Omega_\delta$. Можно указать достаточные условия, обеспечивающие такое неравенство.

Теорема 3. *Пусть выполнены условия теоремы* 7.3. *Тогда существует такая константа* $\Delta_0>0$, *что при любом фиксированном* δ ($\|\delta\|<\Delta_0$) *выполнено неравенство* $\Omega^*\equiv\inf\{\Omega(z)\colon z\in D\}<\Omega_\delta\equiv\lim\limits_{\alpha\to+0}\gamma(\alpha)$.

Доказательство. Предположим, что для некоторой последовательности $\delta_n\in\mathbf{R}^k_+$ ($\|\delta_n\|\to0$, $n\to\infty$) оказывается выполненным равенство $\Omega^*=\Omega_{\delta_n}\equiv\Omega_n$. Тогда по лемме 6.4 и вследствие монотонности по α при каждом фиксированном δ_n функции $\gamma_n(\alpha)\equiv\gamma_{\delta_n}(\alpha)$ оказывается выполненным соотношение $\Omega_n=\gamma_n(+0)\geqslant\gamma_n(\alpha)\geqslant\gamma_n(+\infty)=\Omega^*$ для любого $\alpha>0$. Поэтому $\gamma_n(\alpha)=\Omega^*$ для любого $\alpha>0$.

На основании этого равенства, как в теореме 7.3, получаем $\beta_{\delta_n}(\alpha)=\beta_n(\alpha)=\text{const}$ для всякого $\alpha>0$, откуда следует, что

$$f(\lambda_n)\equiv f(\lambda_{\delta_n})\geqslant f[J^*_{\delta_n}]=\lim_{\alpha\to+0}\beta_n(\alpha)=\lim_{\alpha\to+\infty}\beta_n(\alpha)=\nu_{\delta_n}\equiv\nu_n.$$

Здесь использованы результаты лемм 6.4, 6.6 и теорема 3.1. При доказательстве теоремы 7.3 установлено, что $\nu_n\to\nu_0$ при $n\to\infty$, а в теореме 3.1 показано, что $\lambda_n\equiv\lambda_{\delta_n}\to J^*$. Поэтому, переходя к пределу при $n\to\infty$ в неравенстве $f(\lambda_n)\geqslant\nu_n$, получаем $f(J^*)\geqslant\nu_0$, а это противоречит доказанному в теореме 7.3 неравенству (7.4). Это и доказывает справедливость неравенства $\Omega^*<\Omega_\delta$ для всякого δ ($\|\delta\|\leqslant\Delta_0$).

Алгоритм *обобщенного принципа квазирешений* (о.п.к.) для экстремальных задач использует данные $\{J_\delta, \Psi, f, \delta, \omega_\delta\}$. Он основан на *выборе параметра регуляризации* $\alpha_\delta>0$ как решения уравнения (1) и на использовании следующего правила отбора экстремали задачи (4.2) из их множества $Z^{\alpha_\delta}_\delta$, отвечающего найденному α_δ. Пусть величины $\alpha_{1,2}(\delta)$, $z^{\alpha_{1,2}}$ введены, как в § 7, по найденному α_δ.

Правило отбора 1. Если выполняется неравенство (7.7), то из множества $Z^{\alpha_\delta}_\delta$ выбирается такой элемент z^{α_δ}, для которого справедливо неравенство

$$\Omega(z^{\alpha_\delta})\geqslant\omega_\delta. \tag{2}$$

Если же выполнено (7.9), то выбирается такой элемент z^{α_δ}, что

$$\Omega(z^{\alpha_\delta})\leqslant\omega_\delta. \tag{3}$$

В теоремах 1—3 обоснована возможность выбора параметра регуляризации по обобщенному принципу квазирешений. Возможны при этом два случая. В первом уравнение (1) имеет обычное решение $\alpha_\delta>0$. Тогда

$$\gamma(\alpha_\delta)=\Omega(z^{\alpha_\delta})=\omega_\delta\quad\forall z^{\alpha_\delta}\in Z^{\alpha_\delta}_\delta$$

и для всякого $z^{\alpha_\delta}\in Z_\delta^{\alpha_\delta}$ выполнены условия (2), (3). Поэтому в данном случае правило отбора является излишним. Во втором случае, когда решение уравнения (1) обобщенное, по леммам 6.8, 6.3 число $\alpha_\delta>0$ будет точкой разрыва функции $\gamma(\alpha)$. Более того, согласно лемме 6.5, существуют такие элементы $z_+^{\alpha_\delta}$, $z_-^{\alpha_\delta}$ из множества $Z_\delta^{\alpha_\delta}$, для которых

$$\gamma(\alpha_\delta-0)=\Omega(z_-^{\alpha_\delta})>\omega_\delta, \quad \gamma(\alpha_\delta+0)=\Omega(z_+^{\alpha_\delta})<\omega_\delta.$$

В этом случае правило отбора выделяет некоторый элемент z^{α_δ} из множества $Z_\delta^{\alpha_\delta}$, так как в качестве z^{α_δ} можно взять, например, $z_-^{\alpha_\delta}$, если выполнено (7.7), и $z_+^{\alpha_\delta}$, если выполнено (7.9).

Найденный по алгоритму о.п.к. элемент z^{α_δ} принимается в качестве приближенного решения задачи (2.4) о поиске Ω-оптимальных решений задачи (2.2).

Отметим, что правило отбора 1 отличается от правила отбора 7.1 алгоритма о.п.н. тем, что вместо условий выбора (7.8), (7.10) используются соответственно условия (2), (3).

Будем предполагать, что для каждого $\delta\in\mathbf{R}_+^k$ ($\|\delta\|\leqslant\Delta_0$, $\delta\neq\theta$) параметр регуляризации $\alpha_\delta>0$ и приближенное решение z^{α_δ} задачи (2.4) выбираются в соответствии с алгоритмом о.п.к. Пусть также выполнены предположения 1—3 из § 2 и $f\in\mathcal{F}^m$.

Теорема 4. *Предположим, что для каждого указанного δ число ω_δ удовлетворяет неравенствам $\Omega^*<\omega_\delta<\Omega_\delta$, $\omega_\delta\geqslant\bar\Omega$, причем $\omega_\delta\to\bar\Omega$ при $\Delta\to0$. Тогда, какова бы ни была последовательность $\{\delta_n\}$, сходящаяся к нулю при $n\to\infty$, соответствующая ей последовательность приближенных решений $z_n\equiv z^{\alpha_{\delta_n}}$ будет τ-сходиться к $\bar Z$ при $n\to\infty$ и, кроме того, $\Omega(z_n)\to\bar\Omega$ при $n\to\infty$.*

Доказательство. Достаточно заметить, что, как и в случае алгоритма о.п.к. (см. доказательство теоремы 7.5), семейства $\{z^{\alpha_\delta}\}$, $\{z^{\alpha_{1,2}(\delta)}\}$, фигурирующие в алгоритме о.п.к., удовлетворяют условиям а)—в), д) из § 5. Для этих семейств, кроме того, согласно правилу отбора алгоритма о.п.к., выполнены неравенства (7.7), (2) или (7.9), (3). Такие условия вместе с а)—в), д) входят в лемму 5.11, если в ней положить $\bar\Omega_\delta\equiv\omega_\delta$. По этой лемме будут выполнены соотношения (7.11), (7.12) для семейства z^{α_δ}, получаемого по алгоритму о.п.к. Для завершения доказательства тогда остается использовать следствие 1 из леммы 5.1.

Теорема 4 обосновывает, что обобщенный принцип квазирешений для экстремальных задач является регуляризирующим алгоритмом (см. также [130]).

§ 9. Обобщенный принцип сглаживающего функционала для экстремальных задач

Введем функцию

$$\varepsilon(\alpha)\equiv\varphi(\alpha)-f[\Psi^p(\delta,\gamma(\alpha))+\lambda_\delta]=$$
$$=M^\alpha[z^\alpha]-f[\Psi^p(\delta,\Omega(z^\alpha))+\lambda_\delta]\equiv E(z^\alpha), \qquad (1)$$

определенную при $\alpha>0$ для каждого фиксированного δ ($\delta\in\mathbf{R}_+^k$, $\|\delta\|\leqslant\Delta_0$). Здесь p $(0<p<1)$—заданное число. Будем считать, что выполнены предположения 1—3 из § 2, а также, что $f \in \mathscr{F}^m$ $(m \geqslant 1)$. Тогда из лемм 6.1—6.6, 6.10 нетрудно получить основные свойства функции (1). Одним из этих свойств будет монотонность, которая, как ясно из леммы 6.10, существенно зависит от знака величины Ω^*. В дальнейшем, не ограничивая общности, будем предполагать, что $\Omega^* \geqslant 0$. Добиться этого можно, например, добавляя к ограниченному снизу функционалу $\Omega(z)$ соответствующую константу. Резюмируем свойства функции $\varepsilon(\alpha)$ при фиксированном δ.

Теорема 1. 1. *Если $\Omega^* \geqslant 0$, то функция* (1) *непрерывна при $\alpha > 0$ всюду, за исключением, быть может, не более чем счетного числа точек разрыва первого рода, являющихся точками многозначности этой функции.*

2. *Для всякой точки разрыва $\alpha_0 > 0$ функции* (1) *существуют элементы $z_+^{\alpha_0}$, $z_-^{\alpha_0}$, принадлежащие множеству Z^{α_0} решений задачи* (4.2) *для $\alpha = \alpha_0$, которые удовлетворяют равенствам*

$$\varepsilon(\alpha_0 \pm 0) = E(z_\pm^{\alpha_0}).$$

3. *Если $\Omega^* > 0$, то функция* (1) *монотонно возрастает, причем $\varepsilon(\alpha) \to +\infty$ при $\alpha \to +\infty$. Если $\Omega^* = 0$ и уравнение* (6.8) *имеет решения, среди которых $\alpha_{\min} > 0$ наименьшее, то функция* (1) *монотонно возрастает при $0 < \alpha < \alpha_{\min}$, а при $\alpha \geqslant \alpha_{\min}$ она постоянна: $\varepsilon(\alpha) = \nu_\delta - f[\Psi^p(\delta, \Omega^*) + \lambda_\delta] \equiv \varepsilon_\infty$. Если $\Omega^* = 0$, но уравнение* (6.8) *не имеет решений, то $\varepsilon(\alpha)$ монотонно возрастает и $\varepsilon(\alpha) \to \varepsilon_\infty$ при $\alpha \to +\infty$.*

4. *При $\alpha \to +\infty$ существует предел*

$$\lim_{\alpha \to +0} \varepsilon(\alpha) = f[J_\delta^*] - f[\Psi^p(\delta, \Omega_\delta) + \lambda_\delta] \equiv \varepsilon_0.$$

Рассмотрим теперь уравнение с монотонной при $\Omega^* \geqslant 0$ функцией (1):

$$\varepsilon(\alpha) = 0. \qquad (2)$$

Теорема 2. *Если при фиксированном δ выполнены неравенства $\varepsilon_0 < 0$, $\varepsilon_\infty > 0$, то уравнение* (2) *имеет единственное решение $\alpha_\delta > 0$.*

Существование решения $\alpha_\delta > 0$ следует из леммы 6.7. Если уравнение (2) имеет обобщенное решение, то оно единственно согласно следствию 6.6. Если же решение уравнения (2) обычное, т.е. $\varepsilon(\alpha_\delta) = 0 < \varepsilon_\infty$, то другого решения быть не может вследствие монотонного возрастания функции (1) при $0 < \alpha < \alpha_{\min} \equiv \{\alpha : \varepsilon(\alpha) = \varepsilon_\infty\}$.

Как и в § 7, 8, можно дать достаточные условия справедливости неравенств $\varepsilon_0 < 0$, $\varepsilon_\infty > 0$.

Теорема 3. *Пусть выполнены условия теоремы 7.3. Тогда существует такая константа $\Delta_0 > 0$, что при любом фиксированном δ $(\|\delta\| \leqslant \Delta_0)$ выполнены неравенства $\varepsilon_0^\delta < 0$, $\varepsilon_\infty^\delta > 0$.*

Доказательство этой теоремы проводится путем почти дословного повторения доказательства теоремы 7.3 с заменой там величин $\Psi(\delta, \Omega^*)$, $\Psi(\delta, \Omega_\delta)$ на $\Psi^p(\delta, \Omega^*)$, $\Psi^p(\delta, \Omega_\delta)$.

Можно сформулировать также условия существования и единственности обычного решения уравнения (2).

Теорема 4. *Пусть $\varepsilon_0 < 0$, $\varepsilon_\infty > 0$ и решение экстремальной задачи минимизации сглаживающего функционала* (4.2) *единственно при каж-*

дом $\alpha>0$. *Тогда уравнение* (2) *имеет единственное обычное решение* $\alpha_\delta>0$.

Это почти очевидно следует из теоремы 2 с учетом вытекающей из следствия 6.1 непрерывности функции $\gamma(\alpha)$, а значит, и функции (1).

Обратимся к изложению алгоритма приближенного определения Ω-оптимальных решений экстремальной задачи (2.2), который назовем *обобщенным принципом сглаживающего функционала* (о.п.с.ф.) для экстремальных задач. Будем выбирать параметр регуляризации $\alpha_\delta>0$ по данным $\{J_\delta, \Psi, f, \delta\}$ как решение уравнения (2). В качестве приближения к множеству $\bar{Z}$ Ω-оптимальных решений задачи (2.2) примем элемент z^{α_δ}, находимый по следующему правилу отбора.

Правило отбора 1. Из множества Z^{α_δ} решений задачи (4.2) при $\alpha=\alpha_\delta$ выберем любой элемент z^{α_δ}, для которого выполнено условие

$$E(z^{\alpha_\delta})\geqslant 0. \tag{3}$$

Реализуемость такого алгоритма можно обосновать следующим образом. На основании теорем 2—4 выбор параметра регуляризации по принципу о.п.с.ф. возможен и единствен. Если α_δ обычное решение уравнения (2), то

$$\varepsilon(\alpha_\delta)=E(z^{\alpha_\delta})=0,$$

и поэтому автоматически выполнено условие (3) правила отбора. В случае, если $\alpha_\delta>0$ —обобщенное решение, на основании леммы 6.8 с учетом возрастания $\varepsilon(\alpha)$ можно сделать вывод, что $\varepsilon(\alpha_\delta+0)>0$. Поскольку по теореме 1 найдется элемент $z_+^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого

$$\varepsilon(\alpha_\delta+0)=E(z_+^{\alpha_0})>0,$$

то для него выполнено условие (3), и поэтому можно, например, положить $z^{\alpha_\delta}=z_+^{\alpha_\delta}$.

Алгоритм о.п.с.ф. обладает важной особенностью. Оказывается, что $\alpha_\delta\to 0$ при $\Delta\to 0$. Это свойство может не выполняться для рассмотренных ранее алгоритмов о.п.н. и о.п.к. Соответствующие примеры будут приведены в гл. 2.

Теорема 5. *Предположим, что* $f\in\mathcal{F}^m$ $(m\geqslant 1)$ *и выполнены предположения* 1—3 *из* § 2. *Пусть, кроме того,* $Z^*\cap Z_0=\varnothing$. *Если для всех* $\delta\in\mathbf{R}_+^k$ $(\delta\neq\theta,\ \|\delta\|\leqslant\Delta_0)$ *величина* $\alpha_\delta>0$ *выбрана по обобщенному принципу сглаживающего функционала, то* $\alpha_\delta\to 0$ *при* $\Delta\to 0$.

Доказательство. Предположим, что это не так и для некоторой последовательности $\{\delta_n\}$, сходящейся при $n\to\infty$ к нулю, соответствующая последовательность $\{\alpha_n\}\equiv\{\alpha_{\delta_n}\}$ подчиняется неравенству $\alpha_n\geqslant\alpha_0>0$. Обозначим через z_n какую-либо экстремаль z^{α_n} задачи (4.2), отвечающую значению $\alpha=\alpha_n$, и аналогично будем обозначать через z_n^0 некоторую экстремаль $z_{\delta_n}^{\alpha_0}$ этой задачи при $\alpha=\alpha_0$. Вместо индекса δ_n в формулах будем коротко писать n. Тогда справедлива оценка

$$0\leqslant\Omega^*\leqslant\Omega(z_n)\leqslant\Omega(z_n^0)=\{M^{\alpha_0}[z_n^0]-I_n(z_n^0)\}/\alpha_0\leqslant$$
$$\leqslant\{M^{\alpha_0}[\bar{z}]-\beta(\alpha_0)\}/\alpha_0\leqslant\{\alpha_0\bar{\Omega}+f[\lambda_n+\Psi(\delta_n,\bar{\Omega})]-f[J_n^*]\}/\alpha_0\equiv S_n.$$

Здесь использованы: вытекающие из невозрастания функции $\gamma(\alpha)$ и из неравенства $\alpha_n \geqslant \alpha_0$ соотношения

$$\gamma(\alpha_n) \equiv \Omega(z_n) \leqslant \gamma(\alpha_0) \equiv \Omega(z_n^0),$$

определение сглаживающего функционала (4.1) при $\alpha=\alpha_0$, неравенство (5.2), а также оценка $\beta(\alpha_0) \geqslant f[J_n^*]$, получающаяся с использованием неубывания функции β из леммы 6.4. По теореме 3.1 $\lambda_n \to J^*$ при $n \to \infty$, а по свойствам функции Ψ оказывается, что $\Psi(\delta_n, \bar{\Omega}) \to 0$. Поэтому, учитывая непрерывность функции f и ее ограниченность снизу, можно сделать вывод об ограниченности сверху величин S_n. Таким образом, для всякого n справедлива оценка

$$0 \leqslant \Omega^* \leqslant \Omega(z_n) \leqslant \Omega(z_n^0) \leqslant S_n \leqslant S = \mathrm{const}. \tag{4}$$

Отсюда и из свойств функции Ψ получается соотношение

$$0 \leqslant \Psi(\delta_n, \Omega(z_n)) \leqslant \Psi(\delta_n, \Omega(z_n^0)) \leqslant \Psi(\delta_n, S) \to 0, \quad n \to \infty,$$

из которого ясно, что

$$\Psi(\delta_n, \Omega(z_n)) \to 0, \quad \Psi(\delta_n, \Omega(z_n^0)) \to 0 \tag{5}$$

при $n \to \infty$ для любой экстремали $z_n \in Z^{\alpha_n}$. Используем это и выберем z_n специальным образом: пусть $z_n = z_+^{\alpha_n}$. Тогда, как указано в следствии 6.5,

$$\varepsilon(\alpha_n + 0) = E(z_+^{\alpha_n}) \geqslant 0 \quad \forall n,$$

так что из (1), (5) получается соотношение

$$0 \leqslant \lim_{n\to\infty} E(z_+^{\alpha_n}) = \lim_{n\to\infty} \{\varphi_n(\alpha_n) - f[\Psi^p(\delta_n, \Omega(z_n)) + \lambda_n]\} =$$
$$= \lim_{n\to\infty} \{\varphi_n(\alpha_n) - f(\lambda_n)\}.$$

С другой стороны, выбирая аналогичным образом $z_n = z_-^{\alpha_n}$ и учитывая, что по следствию 6.5

$$\varepsilon(\alpha_n - 0) = E(z_-^{\alpha_n}) \leqslant 0 \quad \forall n,$$

получаем

$$0 \geqslant \lim_{n\to\infty} E(z_-^{\alpha_n}) = \lim_{n\to\infty} \{\varphi_n(\alpha_n) - f[\Psi^p(\delta_n, \Omega(z_n)) + \lambda_n]\} =$$
$$= \lim_{n\to\infty} \{\varphi_n(\alpha_n) - f(\lambda_n)\}.$$

Тем самым

$$\lim_{n\to\infty} \varphi_n(\alpha_n) = \lim_{n\to\infty} f(\lambda_n) = f[J^*].$$

Но на основании неравенства $\varphi(\alpha_0) \leqslant \varphi(\alpha_n)$, вытекающего из не убывания функции φ, и с использованием экстремальности элемента z_n как решения задачи (4.2) с $\alpha=\alpha_n$ можно получить, используя также (4), оценку

$$\varphi_n(\alpha_n) = M_n^{\alpha_n}[z_n] \leqslant M_n^{\alpha_n}[z_n^0] =$$
$$= (\alpha_n - \alpha_0)\Omega(z_n^0) + \varphi_n(\alpha_0) \leqslant S \cdot (\alpha_n - \alpha_0) + \varphi_n(\alpha_n). \tag{6}$$

Переходя здесь к пределу при $n \to \infty$, получаем, что

$$M_n^{\alpha_0}[z_n^0] = \varphi_n(\alpha_0) \to f(J^*). \tag{7}$$

Вернемся теперь к оценке (4). Согласно предположению 3 из § 2, последовательность $\{z_n^0\}$ на основании (4) будет τ-секвенциально компактной в $\Omega_s \subset D$. Значит, существует подпоследовательность $\tilde{z}_k^0 \equiv z_{n_k}^0$ такая, что $\tilde{z}_k^0 \xrightarrow{\tau} z^0 \in D$ при $k \to \infty$. Используя τ-секвенциальную полунепрерывность снизу функционала J, монотонность и непрерывность функции f, условие аппроксимации (2.3), а также соотношения (4)—(6), можно установить неравенство

$$f(J^*) = \inf\{f[J(z)]\colon z \in D\} \leqslant f[J(z^0)] \leqslant \varliminf_{k \to \infty} f[J(\tilde{z}_k^0) - \Psi(\delta_{n_k}, \Omega(\tilde{z}_k^0))] \leqslant$$

$$\leqslant \lim_{k \to \infty} f[J_{n_k}(\tilde{z}_k^0)] = \lim_{k \to \infty} \{\varphi_{n_k}(\alpha_0) - \alpha_0 \Omega(\tilde{z}_k^0)\} \leqslant f(J^*) - \alpha_0 \Omega^* \leqslant f(J^*).$$

Отсюда ясно, что

$$\lim_{k \to \infty} f[J_{n_k}(\tilde{z}_k^0)] \equiv \lim_{k \to \infty} I_{n_k}(\tilde{z}_k^0) = f[J(z^0)] = f(J^*),$$

т. е. $z^0 \in Z^*$. Из (4) с использованием условия аппроксимации (2.3) и монотонности функции f вытекает неравенство

$$0 \leqslant \Omega^* \leqslant \Omega(\tilde{z}_k^0) \leqslant \{\varphi_{n_k}(\alpha_0) - I_{n_k}(z_{n_k}^0)\}/\alpha_0 \leqslant \{\varphi_{n_k}(\alpha_0) - f[J(z_{n_k}^0) -$$

$$- \Psi(\delta_{n_k}, \Omega(z_{n_k}^0))]\}/\alpha_0 \leqslant \{\varphi_{n_k}(\alpha_0) - f[J^* - \Psi(\delta_{n_k}, \Omega(z_{n_k}^0))]\}/\alpha_0. \tag{8}$$

Переходя здесь к пределу при $k \to \infty$ и применяя (5), (7), приходим к соотношению

$$0 \leqslant \Omega^* \leqslant \Omega(z^0) \leqslant \varliminf_{k \to \infty} \Omega(\tilde{z}_k^0) \leqslant \varlimsup_{k \to \infty} \Omega(\tilde{z}_k^0) \leqslant [f(J^*) - f(J^*)]/\alpha_0 = 0,$$

где использована также τ-секвенциальная полунепрерывность снизу функционала Ω. Это означает, что $\Omega(\tilde{z}_k^0) \to \Omega(z^0) = \Omega^* = 0$ при $k \to \infty$, т. е. $z^0 \in Z_0$. Таким образом, установлено, что $z^0 \in Z_0 \cap Z^*$, а это противоречит условию $Z_0 \cap Z^* = \varnothing$. Теорема доказана.

Отметим, что теоремы 1—5 остаются справедливыми и в случае $p = 1$.

Пусть теперь для каждого $\delta \in \mathbf{R}_+^k$ ($\|\delta\| \leqslant \Delta_0$) параметр регуляризации α_δ и элемент $z^{\alpha_\delta} \in Z^{\alpha_\delta}$ найдены по обобщенному принципу сглаживающего функционала. При этом будем считать, что выполнены предположения 1—3 из § 2 и что $f(x) \in \mathscr{F}^m$ для некоторого $m \geqslant 1$.

Теорема 6. *Предположим, что справедливы следующие дополнительные условия*: а) $\Omega^* \geqslant 0$; б) $Z^* \cap Z_0 = \varnothing$; в) *мера аппроксимации* $\Psi(\delta, \Omega)$ *монотонно возрастает по* Ω ($\Omega \geqslant \Omega^*$) *при каждом фиксированном допустимом* δ. *Тогда для всякой последовательности* $\{\delta_n\}$, *сходящейся при* $n \to \infty$ *к нулю, соответствующая последовательность* $z_n \equiv z^{\alpha_{\delta_n}}$ τ-*сходится к множеству* Ω-*оптимальных решений задачи* (2.2). *При этом* $\Omega(z_n) \to \bar{\Omega}$.

Доказательство. Оно базируется на установлении соотношений типа (7.11), (7.12) для выбранных по алгоритму о. п. с. ф. величин α_δ, z^{α_δ} с последующим использованием леммы 5.1 и ее следствия 1.

Пусть $\|\delta\|$ — столь малое число, что $\|\delta\| \leqslant \Delta_0$, $\Psi(\delta, \bar{\Omega}_\delta) < 1$. Здесь величины $\bar{\Omega}_\delta$ выбраны так, чтобы выполнялись условия $\bar{\Omega}_\delta \geqslant \bar{\Omega}$, $\bar{\Omega}_\delta \to \bar{\Omega}$ при $\Delta \to 0$. Достаточно рассмотреть два случая. В первом выполнено неравенство $\Omega(z^{\alpha_\delta}) \leqslant \bar{\Omega}_\delta$. Тогда соотношение (7.12) очевидно, а равенство (7.11) получается из (7.12) с учетом доказанной в теореме 5 сходимости $\alpha_\delta \to 0$ на основании следствия 5.5.

Во втором, более сложном случае выполнено неравенство

$$\Omega(z^{\alpha_\delta}) > \bar{\Omega}_\delta. \tag{9}$$

Тогда по аналогии с неравенством (5.11) можно получить оценку

$$\begin{gathered}0 < \Omega(z^{\alpha_\delta}) - \bar{\Omega}_\delta \leqslant \{f[\lambda_\delta + \Psi(\delta, \bar{\Omega}_\delta)] - f(\lambda_\delta)\}/\alpha_\delta + f'[J_\delta(z^{\alpha_\delta}) + \Psi(\delta, \Omega(z^{\alpha_\delta}))] \times \\ \times \Psi(\delta, \Omega(z^{\alpha_\delta}))/\alpha_\delta = f'(\lambda_\delta^*)\Psi(\delta, \bar{\Omega}_\delta)/\alpha_\delta + f'[J_\delta(z^{\alpha_\delta}) + \Psi(\delta, \Omega(z^{\alpha_\delta}))] \times \\ \times \Psi(\delta, \Omega(z^{\alpha_\delta}))/\alpha_\delta \leqslant \{f'(\lambda_\delta^*) + f'[J_\delta(z^{\alpha_\delta}) + \Psi(\delta, \Omega(z^{\alpha_\delta}))]\} \times \\ \times \{\Psi(\delta, \Omega(z^{\alpha_\delta}))/\alpha_\delta\},\end{gathered} \tag{10}$$

где использованы формула конечных приращений, а также вытекающее из (9) и монотонности функции Ψ неравенство

$$\Psi(\delta, \bar{\Omega}_\delta) < \Psi(\delta, \Omega(z^{\alpha_\delta})).$$

Величина λ_δ^* ограничена, так как удовлетворяет неравенству $J^* \leqslant \lambda_\delta \leqslant \lambda_\delta^* \leqslant \lambda_\delta + \Psi(\delta, \bar{\Omega}_\delta) < \lambda_\delta + 1 \leqslant \lambda_{\max} = \text{const}$ (см. теорему 3.1). Если теперь будет показано, что при $\Delta \to 0$

$$\Psi(\delta, \Omega(z^{\alpha_\delta}))/\alpha_\delta \to 0, \tag{11}$$

то вследствие сходимости $\alpha_\delta \to 0$ окажется, что и $\Psi(\delta, \Omega(z^{\alpha_\delta})) \to 0$ при $\Delta \to 0$. Тогда, согласно следствию 5.5, имеет место сходимость $J_\delta(z^{\alpha_\delta}) \to J^*$, из которой получается (7.11). Из этой же сходимости, а также из (11), примененных к неравенству (10), можно получить, учитывая непрерывность производной f', что $\Omega(z^{\alpha_\delta}) - \bar{\Omega}_\delta \to 0$ при $\Delta \to 0$. Поэтому и $\Omega(z^{\alpha_\delta}) \to \bar{\Omega}$, так что неравенство (7.12) также выполнено.

Таким образом, остается доказать соотношение (11). Справедлива оценка

$$\begin{gathered}f[\Psi^p(\delta, \bar{\Omega}_\delta) + \lambda_\delta] \leqslant f[\Psi^p(\delta, \Omega(z^{\alpha_\delta})) + \lambda_\delta] \leqslant \varphi(\alpha_\delta) \leqslant M^{\alpha_\delta}[\bar{z}] \leqslant \\ \leqslant \alpha_\delta \bar{\Omega} + f[\Psi(\delta, \bar{\Omega}) + \lambda_\delta] \leqslant \alpha_\delta \bar{\Omega}_\delta + f[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta].\end{gathered} \tag{12}$$

Здесь первое и последнее неравенства получены, исходя из возрастания функций f, Ψ, а также с учетом неравенств (9) и $\bar{\Omega}_\delta \geqslant \bar{\Omega}$. Второе неравенство представляет собой следствие из условия (3) выбора элемента z^{α_δ}. Из экстремальности этого элемента как решения задачи (4.2) вытекает третье неравенство, а четвертое неравенство совпадает с (5.2).

Из (12), в частности, можно получить оценку для параметра регуляризации:

$$\begin{gathered}\alpha_\delta \geqslant \{f[\Psi^p(\delta, \bar{\Omega}_\delta) + \lambda_\delta] - f[\Psi(\delta, \bar{\Omega}_\delta) + \lambda_\delta]\}\bar{\Omega}_\delta^{-1} = \\ = \bar{\Omega}_\delta^{-1} f'(\tilde{\lambda})[\Psi^p(\delta, \bar{\Omega}_\delta) - \Psi(\delta, \bar{\Omega}_\delta)] \geqslant \bar{\Omega}_\delta^{-1} s_\delta^m [\Psi^p(\delta, \bar{\Omega}_\delta) - \Psi(\delta, \bar{\Omega}_\delta)],\end{gathered} \tag{13}$$

где равенство получено на основании формулы конечных приращений и $\tilde{\lambda}$ подчинена неравенству

$$\Psi^p(\delta, \bar{\Omega}_\delta)+\lambda_\delta \geqslant \tilde{\lambda} \geqslant \Psi(\delta, \bar{\Omega}_\delta)+\lambda_\delta \tag{14}$$

Число s_δ^m, оценивающее снизу значение $f'(\tilde{\lambda})$, меняется в зависимости от предположений относительно функции f. Если $f \in \mathscr{F}^1$ $(m=1)$, то можно положить $f'(\tilde{\lambda}) \geqslant s_\delta^1 \equiv \varkappa_0 > 0$ (см. § 4). Если же $f \in \mathscr{F}^m$ для $m>1$, то в силу возрастания производной f' из оценки (14) получим

$$f'(\tilde{\lambda}) \geqslant s_\delta^m = f'[\Psi(\delta, \bar{\Omega}_\delta)+\lambda_\delta] > f'[\Psi(\delta, \Omega^*)+\lambda_0] \geqslant 0.$$

Строгое неравенство здесь получено следующим образом. Из условия $Z_0 \cap Z^* = \varnothing$ вытекает, что $\bar{\Omega} = \inf\{\Omega(z)\colon z \in Z^*\} > \inf\{\Omega(z)\colon z \in Z_0\} = \inf\{\Omega(z)\colon z \in D\} \equiv \Omega^*$, т. е. $\bar{\Omega}_\delta \geqslant \bar{\Omega} > \Omega^*$. Далее используется монотонное возрастание функции f' и меры аппроксимации $\Psi(\delta, \Omega)$ по аргументу Ω.

Таким образом. $s_\delta^m > 0$. Далее, из (12) ясно, что

$$0 \leqslant \bar{\chi}(\delta) \equiv \{f[\Psi^p(\delta, \Omega(z^{\alpha_\delta}))+\lambda_\delta] - f(\lambda_\delta)\}/\alpha_\delta^p \leqslant$$
$$\leqslant \{\varphi(\alpha_\delta) - f(\lambda_\delta)\}/\alpha_\delta^p = \bar{\Delta}_1(\delta) + \bar{\Delta}_2(\delta), \tag{15}$$

где

$$\bar{\Delta}_1(\delta) \equiv \{\varphi(\alpha_\delta) - f[\Psi(\delta, \bar{\Omega}_\delta)+\lambda_\delta]\}/\alpha_\delta^p,$$
$$\bar{\Delta}_2(\delta) \equiv \{f[\Psi(\delta, \bar{\Omega}_\delta)+\lambda_\delta] - f(\lambda_\delta)\}/\alpha_\delta^p.$$

Из (12), кроме того, получается оценка

$$0 \leqslant \bar{\Delta}_1(\delta) \leqslant \alpha_\delta^{1-p}\bar{\Omega}_\delta. \tag{16}$$

Для величины $\bar{\Delta}_2(\delta)$, используя (13) и формулу конечных приращений, получаем

$$0 \leqslant \bar{\Delta}_2(\delta) = \{f'[\lambda_\delta + \theta_0\Psi(\delta, \bar{\Omega}_\delta)]\,\Psi(\delta, \bar{\Omega}_\delta)/\alpha_\delta\}\,\alpha_\delta^{1-p} \leqslant$$
$$\leqslant \{f'[\lambda_\delta + \theta_0\Psi(\delta, \bar{\Omega}_\delta)]/s_\delta^m\} \cdot \bar{\Omega}_\delta\Psi(\delta, \bar{\Omega}_\delta)\,\alpha_\delta^{1-p}/[\Psi^p(\delta, \bar{\Omega}_\delta) - \Psi(\delta, \bar{\Omega}_\delta)] \equiv$$
$$\equiv M_\delta^m\bar{\Omega}_\delta\Psi^{1-p}(\delta, \bar{\Omega}_\delta)\,\alpha_\delta^{1-p}/[1-\Psi^{1-p}(\delta, \bar{\Omega}_\delta)], \quad 0<\theta_0<1. \tag{17}$$

Величина

$$M_\delta^m \equiv f'[\lambda_\delta + \theta_0\Psi(\delta, \bar{\Omega}_\delta)]/s_\delta^m$$

оценивается сверху при $m=1$ следующим образом:

$$M_\delta^1 \leqslant \sup_\delta\{f'[\lambda_\delta + \theta_0\Psi(\delta, \bar{\Omega}_\delta)]\,\varkappa_0^{-1}\colon \|\delta\| \leqslant \Delta_0\} \equiv M_0,$$

а при $m>1$ вследствие возрастания функции f' величина M_δ^m оценивается сверху числом 1:

$$M_\delta^m \leqslant f'[\lambda_\delta + \Psi(\delta, \bar{\Omega}_\delta)]/s_\delta^m = 1.$$

Таким образом, из (16), (17), переходя к пределу при $\Delta \to 0$, учитывая установленную ограниченность чисел M_δ^m, а также сходимости $\alpha_\delta \to 0$, $\Psi(\delta, \bar{\Omega}_\delta) \to 0$ при $\Delta \to 0$, получаем $\bar{\Delta}_{1,2}(\delta) \to 0$. Тогда из (15) ясно, что $\bar{\chi}(\delta) \to 0$. Установим следствия этой сходимости при различных $m \geqslant 1$.

Пусть $m=1$. В этом случае, применяя в (15) формулу конечных приращений, получаем с учетом определения функции из $\mathscr{F}^1$

$$\bar{\chi}(\delta)=f'(\hat{\lambda}_\delta)\Psi^p(\delta,\Omega(z^{\alpha_\delta}))/\alpha_\delta^p\geqslant\varkappa_0\left[\Psi(\delta,\Omega(z^{\alpha_\delta}))/\alpha_\delta\right]^p>0,$$

где $\lambda_\delta\leqslant\hat{\lambda}_\delta$. Переходя в этом неравенстве к пределу при $\Delta\to0$, легко теперь получить (11). Аналогичные рассуждения могут быть использованы для получения соотношения (11) и в случае, когда $m>1$, $f'(J^*)>0$, с заменой $\varkappa_0$ на $f'(J^*)$.

Пусть теперь $m>1$, $f'(J^*)=0$. В этом случае не будем доказывать (11), а непосредственно установим соотношения (7.11), (7.12). Будем при этом использовать схему, предложенную для рассмотрения аналогичного случая в лемме 5.10. В этой лемме, в частности, установлено неравенство (5.29), принимающее вид

$$0<\Omega(z^{\alpha_\delta})-\bar{\Omega}_\delta\leqslant\sum_{k=1}^{m-1}\frac{1}{k!\,\alpha_\delta}f^{(k)}(J^*)\Psi^k(\delta,\bar{\Omega}_\delta)+\frac{1}{m!}f^{(m)}\left[J^*+\theta\Psi(\delta,\bar{\Omega}_\delta)\right]\times$$

$$\times\frac{\Psi^m(\delta,\bar{\Omega}_\delta)}{\alpha_\delta}\leqslant\sum_{k=1}^{m-1}\frac{1}{k!}f^{(k)}(J^*)\,\frac{\Psi^k(\delta,\Omega(z^{\alpha_\delta}))}{\alpha_\delta}+N_1\,\frac{\Psi^m(\delta,\Omega(z^{\alpha_\delta}))}{\alpha_\delta}\equiv$$

$$\equiv\tilde{\Delta}_1(\delta)+N_1\tilde{\Delta}_2(\delta),\quad 0<\theta<1.\tag{18}$$

Кроме того, из (15) следует

$$\bar{\chi}(\delta)=\sum_{k=1}^{m-1}\frac{1}{k!}f^{(k)}(\lambda_\delta)\left[\frac{\Psi^k(\delta,\Omega(z^{\alpha_\delta}))}{\alpha_\delta}\right]^p+\frac{1}{m!}f^{(m)}\left[\lambda_\delta+\theta_1\Psi^p(\delta,\Omega(z^{\alpha_\delta}))\right]\times$$

$$\times\left[\frac{\Psi^m(\delta,\Omega(z^{\alpha_\delta}))}{\alpha_\delta}\right]^p\geqslant\tilde{\Delta}_1(\delta)+\frac{1}{m!}f^{(m)}(\lambda_\delta+\theta_1\Psi^p(\delta,\Omega(z^{\alpha_\delta})))\cdot\tilde{\Delta}_2(\delta)\geqslant$$

$$\geqslant\tilde{\Delta}_1(\delta)+\frac{\varkappa_0}{m!}\tilde{\Delta}_2(\delta)\geqslant0,\quad 0<\theta_1<1.\tag{19}$$

Соотношение (19) получено по формуле Тейлора. В нем также учтены неравенства $f^{(k)}(\lambda_\delta)\geqslant f^{(k)}(J^*)$, $f^{(m)}(x)\geqslant\varkappa_0$ $(k=1,\ldots,m-1)$, вытекающие из монотонного возрастания производных $f^{(k)}$ (см. § 4), неравенства $\lambda_\delta\geqslant J^*$ и определения класса функций $\mathscr{F}^m$. Из (19) ясно, что в силу сходимости $\bar{\chi}(\delta)\to0$ при $\Delta\to0$ неотрицательные величины $\tilde{\Delta}_{1,2}(\delta)$ также стремятся к нулю. Но тогда, согласно (18), будет выполнено соотношение $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$ при $\Delta\to0$, т. е. будет справедливо (7.12). Равенство (7.11) получается из доказанного (7.12), сходимости $\alpha_\delta\to0$ и следствия 5.5. Теорема доказана.

Теорема 6 доказывает, что алгоритм о.п.с.ф. представляет собой регуляризирующий алгоритм для решения некорректных экстремальных задач. Необходимо подчеркнуть, что при $p=1$ алгоритм о.п.с.ф. может и не обеспечивать сходимость приближений z^{α_δ} к множеству $\bar{Z}$ Ω-оптимальных решений задачи (2.2). Соответствующий пример будет приведен в § 2.2.

Замечание. Алгоритм о.п.с.ф. для экстремальных задач предложен в работе [130]. В частном случае равномерной аппроксимации функционала $J(z)$ приближенным функционалом $J_\delta(z)$ на множестве D, когда можно положить $\Psi(\delta, \Omega)\equiv\Psi(\delta)$, алгоритм о.п.с.ф. переходит в изучавшийся в работах [151, 142] принцип сглаживающего функционала.

§ 10. Некоторые дополнительные свойства обобщенных принципов невязки, квазирешений и сглаживающего функционала

В § 7—9 показано, что если значение параметра регуляризации α_δ, выбираемое по сформулированным там методам, положительно, то соответствующие алгоритмы построения приближенного решения z^{α_δ} дают устойчивый метод решения основной задачи второго типа (см. § 4). Возникает вопрос: что будет в особом случае, если обобщенные принципы невязки, квазирешений или сглаживающего функционала дают $\alpha_\delta=0$? В данном параграфе будет указан способ выбора устойчивого приближения к множеству Ω-оптимальных решений задачи (2.2) и в этом особом случае.

Будем обозначать α_δ(о.п.н.), α_δ(о.п.к.), α_δ(о.п.с.ф.) параметры регуляризации, выбираемые по соответствующему алгоритму. Считаем также, что выполнены предположения 1—3 из § 2, а функция $f(x)$ монотонно возрастает при $x\geqslant x_0$.

Лемма 1. 1. *Равенство* α_δ(о.п.н.)$=0$ *возможно в том и только том случае, если* α_δ(о.п.с.ф.)$=0$.

2. *Если дополнительно к обычным свойствам функции* Ψ *она монотонно возрастает по второму аргументу для данного* δ ($\|\delta\|\leqslant\Delta_0$, $\delta\neq\theta$), *то из равенства* α_δ(о.п.н.)$=0$ (*или* α_δ(о.п.с.ф.)$=0$) *следует, что* α_δ(о.п.к.)$=0$.

Доказательство. Рассмотрим п. 1. Согласно следствию 6.4, уравнения (7.3), (9.2) имеют нулевое решение тогда и только тогда, когда

$$\rho_0^\delta=\lim_{\alpha\to+0}\rho(\alpha)\equiv f[J_\delta^*]-f[\Psi(\delta, \Omega_\delta)+\lambda_\delta]\geqslant 0, \tag{1}$$

$$\varepsilon_0^\delta=\lim_{\alpha\to+0}\varepsilon(\alpha)\equiv f[J_\delta^*]-f[\Psi^p(\delta, \Omega_\delta)+\lambda_\delta]\geqslant 0, \tag{2}$$

$$0<p<1.$$

Выражения для этих пределов взяты из результатов теорем 7.1, 9.1.

Предположим теперь, что α_δ(о.п.н.)$=0$, т. е. выполнено первое из соотношений (1). Тогда вследствие монотонного возрастания функции f, а также с учетом неравенств $\Psi\geqslant 0$, $\lambda_\delta\geqslant J_\delta^*$ (см. теорему 3.1) получим

$$J_\delta^*\geqslant\Psi(\delta, \Omega_\delta)+\lambda_\delta\geqslant\Psi(\delta, \Omega_\delta)+J_\delta^*\geqslant J_\delta^*. \tag{3}$$

Отсюда ясно, что

$$\Psi(\delta, \Omega_\delta)=0, \quad J_\delta^*=\lambda_\delta. \tag{4}$$

Эти равенства в свою очередь приводят к выполнению (2), а, значит, по следствию 6.4 — к выводу, что α_δ(о.п.с.ф.)$=0$. Аналогично доказывается, что из равенства α_δ(о.п.с.ф.)$=0$ следует α_δ(о.п.к.)$=0$.

Используем теперь равенства (4), представляющие собой следствие условий α_δ(о.п.н.)$=0$, α_δ(о.п.с.ф.)$=0$, для доказательства п. 2 леммы. Поскольку $\Omega_\delta \geqslant \Omega^*$, то вследствие монотонного возрастания функции Ψ получим $0=\Psi(\delta, \Omega_\delta) \geqslant \Psi(\delta, \Omega^*)$, т. е. $\Psi(\delta, \Omega_\delta)=\Psi(\delta, \Omega^*)=0$. Это возможно лишь при выполнении равенства $\Omega_\delta=\Omega^*$. В свою очередь последнее равенство с учетом леммы 6.4 означает, что $\gamma(\alpha)=\Omega^*$ для любого $\alpha>0$. Таким образом, $\gamma(+0)=\Omega^*<\omega_\delta$ (ср. условие теоремы 8.1). Это по следствию 6.4 и дает равенство α_δ(о.п.к.)$=0$.

Следствие 1. *Из возрастания функции f, возрастания функции Ψ по второму аргументу для данного δ ($\delta \neq \theta$, $\|\delta\| \leqslant \Delta_0$) и из равенства α_δ(о.п.н.)$=0$ следует, что $\Omega(z^\alpha)=\gamma(\alpha)=\Omega^*$ для всякого $\alpha>0$.*

На простых примерах нетрудно убедиться, что из равенства α_δ(о.п.к.)$=0$ могут и не следовать равенства α_δ(о.п.н.)$=0$ или α_δ(о.п.с.ф.)$=0$ (см. пример 2.2.1).

Пусть значение параметра регуляризации α_δ, выбранное по одному из алгоритмов § 7—9, оказывается нулевым. Формально при этом задача минимизации сглаживающего функционала (4.2) принимает следующий вид: найти такой элемент $z_\delta^0 \in D$, для которого

$$J_\delta(z_\delta^0)=\inf\{J_\delta(z)\colon\ z \in D\} \equiv J_\delta^*. \qquad (5)$$

При этом остается пока открытым вопрос о разрешимости этой задачи и свойствах ее решений. Условия, наложенные на функционал J_δ в § 2, вообще говоря, не гарантируют существования решений задачи (5), и только дополнительное условие $\alpha_\delta=0$ обеспечивает ее разрешимость.

Теорема 1. *Предположим, что мера аппроксимации Ψ при данном $\delta \in \mathbf{R}_+^k$ ($\delta \neq \theta$) обладает дополнительным свойством: она монотонно возрастает по второму аргументу. Тогда из равенства α_δ(о.п.н.)$=0$ (или α_δ(о.п.с.ф.)$=0$) следует, что задача* (5) *разрешима, причем выполнено равенство*

$$\inf\{J_\delta(z)\colon\ z \in D\}=\inf\{J_\delta(z)\colon\ z \in Z_0\}=J_\delta(z_\delta^0). \qquad (6)$$

Доказательство. Следствие 1 дает равенство

$$\Omega(z^\alpha)=\gamma(\alpha)=\Omega^* \quad \forall \alpha>0. \qquad (7)$$

Согласно предположению 2 из § 2 это обеспечивает τ-секвенциальную компактность множества $\{z^\alpha\}$. В частности, если $\{\alpha_n\}$ — сходящаяся к нулю последовательность, то из соответствующей последовательности $\{z_n\} \equiv \{z^{\alpha_n}\}$ можно выбрать τ-сходящуюся подпоследовательность $\{z_{n_k}\}$: $z_{n_k} \xrightarrow{\tau} z_\delta^0 \in D$. Используя τ-секвенциальную полунепрерывность снизу функционала J_δ и лемму 6.4, можно тогда получить

$$J_\delta^*=\lim_{\alpha \to +0} J_\delta(z^\alpha)=\varliminf_{k \to \infty} J_\delta(z_{n_k}) \geqslant J_\delta(z_\delta^0) \geqslant J_\delta^*=\inf\{J_\delta(z)\colon\ z \in D\},$$

откуда очевидно равенство $J_\delta(z_\delta^0)=J_\delta^*$. Это равенство означает, что элемент z_δ^0 является решением задачи (5). Аналогично, используя τ-секвенциальную полунепрерывность снизу функционала Ω, можно получить с учетом (7), что

$$\Omega(z_\delta^0) \leqslant \varliminf_{k\to\infty} \Omega(z_{n_k}) = \lim_{\alpha\to+0} \Omega(z^\alpha) = \Omega^* = \inf\{\Omega(z):\ z\in D\}. \tag{8}$$

Тем самым $\Omega(z_\delta^0)=\Omega^*$, $z_\delta^0\in Z_0$. Это доказывает равенство (6).

Лемма 2. *Пусть условие теоремы* 1 *на функцию* Ψ *и равенство* α_δ(о.п.н.)$=0$ (*или* α_δ(о.п.с.ф.)$=0$) *выполнено для каждого* δ $(0<\|\delta\|\leqslant\Delta_0)$. *Тогда* $Z_0\cap Z^*\neq\varnothing$.

Справедливость этой леммы непосредственно вытекает из теорем 7.2, 7.3.

Таким образом, параметр регуляризации, выбираемый по о.п.н. или о.п.с.ф., равен нулю для всех достаточно малых Δ лишь в случае, если Ω-оптимальные решения задачи (2.2) реализуют глобальный минимум функционала Ω на множестве D.

Теорема 2. *Если* α_δ(о.п.к.)$=0$, *то задача* (5) *разрешима и для всякого ее решения справедливо неравенство* $\Omega(z_\delta^0)\leqslant\omega_\delta$.

Доказательство. Согласно следствию 6.4, равенство α_δ(о.п.к.)$=0$ возможно, если лишь $\gamma(+0)=\Omega_\delta\leqslant\omega_\delta$. Тогда по леммам 6.1, 6.4 имеем $\gamma(\alpha)\leqslant\gamma(+0)\leqslant\omega_\delta$ для всякого $\alpha>0$. Это неравенство является аналогом используемого при доказательстве теоремы 1 неравенства (7) и обеспечивает τ-секвенциальную компактность множества $\{z^\alpha\}$. Дальнейшие рассуждения проводятся, как в теореме 1, с заменой в (8) величины Ω^* на ω_δ.

Результаты теорем 1, 2 и лемм 1, 2 позволяют сформулировать правило выбора приближенного решения в особых случаях $\alpha_\delta=0$ для алгоритмов о.п.н., о.п.с.ф. и о.п.к.

Правило 1. Если α_δ(о.п.н.)$=\alpha_\delta$(о.п.с.ф.)$=0$, то в качестве приближения к множеству $\bar{Z}$ при данном δ принимаем любое решение экстремальной задачи: найти элемент z_δ^0, для которого

$$J_\delta(z_\delta^0)=\inf\{J_\delta(z):\ z\in Z_0\}. \tag{9}$$

Если α_δ(о.п.к.)$=0$, то за приближенное решение можно принять любое решение задачи (5).

Правило 1, по существу, является дополнением правил отбора, данных в § 7—9.

Поэтому будем условно называть элементы, полученные из (9) при α_δ(о.п.н.)$=0$ (или при α_δ(о.п.с.ф.)$=0$), *приближениями алгоритмов о.п.н.* или *о.п.с.ф.* при $\alpha_\delta=0$. Аналогично будем называть решения задачи (5) при α_δ(о.п.к.)$=0$ *приближениями алгоритма о.п.к.*

В ряде случаев задача (9) выглядит особенно просто. Так, например, если функционал Ω имеет единственную точку глобального минимума на D, т. е. множество Z_0 состоит из одного элемента z_0, то этот элемент и будет решением задачи (9).

Предположим теперь, что имеется некоторая последовательность $\{\delta_n\}$, сходящаяся к нулю при $n\to\infty$. Будем считать, что при каждом таком δ_n какой-нибудь из рассматриваемых в § 7—9 алгоритмов формально дает $\alpha_{\delta_n}\equiv\alpha_n=0$, и поэтому приближенное решение $z_n\equiv z_{\delta_n}^0$ выбирается по правилу 1. Будем также предполагать, что в случае алгоритмов о.п.н. и о.п.с.ф. для функции Ψ выполнено условие теоремы 1 при всех δ $(0<\|\delta\|\leqslant\Delta_0)$.

Теорема 3. *При* $n\to\infty$ *имеют место сходимости* $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$.

Доказательство. Если приближения z_n получены из (9) по алгоритмам о.п.н. или о.п.с.ф., то вследствие леммы 2 справедливо равенство $\bar{\Omega}=\Omega^*$ и, значит, $\Omega(z_n)\leqslant\bar{\Omega}$ (см. (8)). С другой стороны, для элементов z_n, полученных по алгоритму о.п.к. из (5), по теореме 2 выполнено неравенство $\Omega(z_n)\leqslant\omega_{\delta_n}$. Поэтому $\Omega(z_n)\leqslant\Omega_n\equiv\max\{\bar{\Omega}, \omega_{\delta_n}\}\leqslant K=\text{const}$. Отсюда ясно, что

$$\varlimsup_{n\to\infty}\Omega(z_n)\leqslant\lim_{n\to\infty}\Omega_n=\bar{\Omega}. \tag{10}$$

Далее, используя условие аппроксимации (2.3) и свойства функции Ψ, можно найти, что

$$J^*=\inf\{J(z):\ z\in D\}\leqslant J(z_n)\leqslant J_{\delta_n}(z_n)+\Psi(\delta_n, \Omega(z_n))\leqslant J^*_{\delta_n}+\Psi(\delta_n, \Omega_n)\leqslant$$
$$\leqslant J^*_{\delta_n}+\Psi(\delta_n, K),$$
$$J^*_{\delta_n}=J_{\delta_n}(z_n)=\inf\{J_{\delta_n}(z):\ z\in D\}\leqslant J_{\delta_n}(\bar{z})\leqslant J(\bar{z})+\Psi(\delta_n, \bar{\Omega})=J^*+\Psi(\delta_n, \bar{\Omega}).$$

Здесь применены также равенства (5), (6). Из двух последних оценок находим

$$J^*-\Psi(\delta_n, K)\leqslant J^*_{\delta_n}=J_{\delta_n}(z_n)\leqslant J^*+\Psi(\delta_n, \bar{\Omega}),$$

откуда вследствие предельного перехода с учетом свойств функции Ψ можно получить равенство

$$\lim_{n\to\infty}J_{\delta_n}(z_n)=J^*. \tag{11}$$

Доказанные соотношения (10), (11) представляют собой условия леммы 5.1, по которой и получается доказательство теоремы 3.

Теорема 2 показывает, что алгоритмы о.п.н., о.п.к. и о.п.с.ф. остаются регуляризующими и в случае, когда они дают нулевое значение пераметра регуляризации. При этом в случае алгоритма о.п.к. задача поиска Ω-оптимальных решений задачи (2.2) оказывается по теореме 2 и определению 1.2 устойчивой в рассматриваемом классе приближенных функционалов.

§ 11. Частные случаи условий применимости алгоритмов о.п.н., о.п.к., о.п.с.ф.

В § 2, 4 сформулирован комплекс условий на функционалы J, J_δ, Ω, множество D и функции f, Ψ, входящие в постановку задачи нахождения Ω-оптимальных решений экстремальной задачи (2.2) и в алгоритмы ее решения. Эти условия носят общий характер и выполняются для многих экстремальных задач (2.2), (2.4). Рассмотрим некоторые типичные частные случаи, когда условия постановки задачи выполнены. Начнем с анализа условия аппроксимации (2.3).

1. Пусть Z — линейное нормированное пространство. Тогда обычным способом выделения класса допустимых приближенных функционалов задачи (2.2) является задание условия аппроксимации вида

$$|J_\delta(z)-J(z)|\leqslant\Psi_0(\delta,\ \|z\|)\quad \forall z\in D\subset Z, \tag{1}$$

где Ψ_0 удовлетворяет предположениям на меру аппроксимации, перечисленным в § 2. Условие (1) формально отличается от (2.3), если $\Omega(z)\not\equiv\|z\|$. Тем не менее условие (2.3) будет выполнено, если $\Omega(z)\geqslant g(\|z\|)$ для любого $z\in D$ и если функция $g(x)$ удовлетворяет условию А: $g(x)$ определена, непрерывна и монотонно возрастает при $0\leqslant x$, $g(+\infty)=+\infty$. Действительно, при выполнении условия А существует возрастающая и непрерывная в своей области определения обратная функция g^{-1} и $\|z\|=g^{-1}[g(\|z\|)]$. Поэтому

$$\Psi_0(\delta,\ \|z\|)=\Psi_0(\delta,\ g^{-1}[g(\|z\|)])\leqslant\Psi_0(\delta,\ g^{-1}[\Omega(z)])\equiv$$
$$\equiv\Psi_1(\delta,\ \Omega(z)), \tag{2}$$

где использованы свойства монотонного возрастания функции g^{-1} и монотонного неубывания по второму аргументу функции $\Psi_0(\delta,\ x)$ $(x\geqslant 0)$. Из (1), (2) ясно, что условие аппроксимации (2.3) выполнено, причем в качестве меры аппроксимации можно взять $\Psi_1(\delta,\Omega)$. Для Ψ_1 выполнены предположения 3, а) — в) из § 2.

Весьма часто на практике требуется найти *нормальное решение* задачи (2.2), т. е. рассмотреть вместо общей (2.4) экстремальную задачу

$$\operatorname{Arg\,inf}\{\|z\|:\ z\in Z^*\}.$$

Тогда можно принять $\Omega(z)=g(\|z\|)$, где $g(x)$ удовлетворяет условию А и выбирается исходя из специфики задачи. Тогда снова условие (2.3) выполнено $\Psi=\Psi_1(\delta,\ \Omega)$. Типичными примерами функционалов Ω, используемых для поиска нормальных решений, являются следующие: $\Omega(z)=\|z\|^q$ $(q=\text{const}>0)$ (см. [190—192, 202, 88] и др.); $\Omega(z)=e^{\|z\|}$, $\Omega(z)=a_1^2\|z\|+\ldots+a_N^2\|z\|^N$, где константы a_k $(k=1,\ \ldots,\ N)$ подчинены условию $a_1^2+\ldots+a_N^2\neq 0$.

В некоторых случаях [202] требуется найти *обобщенное нормальное решение* задачи (2.2) относительно элемента $z_0\in D$. Тогда полагают $\Omega(z)=g_0(\|z-z_0\|)$, где $g_0(x)=g(x+\|z_0\|)$, а g удовлетворяет условию А. Легко видеть, что $\Omega(z)=g(\|z-z_0\|+\|z_0\|)\geqslant g(\|z\|)$, и, как показано выше, из (1) следует условие (2.3) с мерой аппроксимации $\Psi=\Psi_1(\delta,\ \Omega)$. Если в дополнении к условию А функция $g(x)$ удовлетворяет также неравенству

$$g(a+b)\leqslant\gamma[g(a)+g(b)]\quad \forall a,b\geqslant 0,\quad \gamma=\text{const}>0, \tag{3}$$

то можно принять $\Omega(z)=g(\|z-z_0\|)$. В этом случае по аналогии с (2) можно получить оценку

$$\Psi_0(\delta,\ \|z\|)=\Psi_0(\delta,\ g^{-1}(g(\|z\|)))\leqslant\Psi_0(\delta,\ g^{-1}[\gamma g(\|z-z_0\|)+\gamma g(\|z_0\|)])\equiv$$
$$\equiv\Psi_2(\delta,\ \Omega(z)). \tag{4}$$

Из (4) понятно, что условие (2.3) выполняется с мерой аппроксимации $\Psi=\Psi_2(\delta,\Omega)$, причем для Ψ_2 вследствие свойств монотонности и непрерывности функций Ψ_0, g будут выполнены предположения 3, а) — в) из § 2.

Пример 1. Предположим, что допустимые приближенные функционалы задачи (2.2) удовлетворяют условию (1), а в качестве $g(\|z\|)$ взята $\|z\|^q$ $(q \geqslant 1)$. Построим функционал Ω и меру аппроксимации Ψ для поиска по схеме из § 4 обобщенного нормального решения задачи (2.2) относительно элемента z_0. Как только что было сказано, можно положить

$$\Omega(z) = g(\|z - z_0\| + \|z_0\|) = [\|z - z_0\| + \|z_0\|]^q, \quad \Psi = \Psi_0(\delta, \Omega^{1/q}).$$

Возможен и другой подход, учитывающий неравенство типа (3) $(a+b)^q \leqslant 2^{q-1}(a^q + b^q)$. Тогда из (4) получим

$$\Psi = \Psi_2(\delta, \Omega) = \Psi_0(\delta, 2^{(q-1)/q}[\Omega(z) + \Omega(0)]^{1/q}), \quad \Omega(z) = \|z - z_0\|^q.$$

Если в задаче (2.2) множество D — выпуклое подмножество линейного нормированного пространства Z, то можно рассмотреть задачу поиска Ω-оптимальных решений для часто встречающихся в приложениях (см. [32, 202] и др.) строго равномерно выпуклых на D функционалов $\Omega(z)$.

Из определения 0.3.22 следует, что для таких функционалов существует монотонно возрастающая функция $v(x)$, определенная при $x \geqslant 0$, для которой $v(0) = 0$ и для каждой пары элементов z_1, $z_2 \in D$ выполнена оценка

$$v(\|z_1 - z_2\|) \leqslant \Omega(z_1) + \Omega(z_2) - 2\Omega((z_1 + z_2)/2). \tag{5}$$

Пусть функция $v(x)$ для рассматриваемого функционала Ω оказывается непрерывной, а множество D содержит нуль-элемент пространства Z, причем $\Omega(0) = \Omega^* = \inf\{\Omega(z): z \in D\}$. Тогда из (5) следует оценка $v(\|z\|) \leqslant \Omega(z) - \Omega^*$, из которой при выполнении условия (1) следует условие аппроксимации (2.3) с мерой аппроксимации $\Psi(\delta, \Omega(z)) = \Psi_0(\delta, v^{-1}(\Omega(z) - \Omega^*))$. В частном случае пространство Z может здесь быть гильбертовым, а функционал Ω — сильно выпуклым (см. § 0.3).

В заключение рассмотрения условий аппроксимации приведем пример меры Ψ_0 из (1):

$$\Psi_0(\delta, \|z\|) = \sum_{i=1}^{k} \delta_i \|z\|^{i-1}, \quad \delta = (\delta_1, \ldots, \delta_k) \in \mathbf{R}_+^k, \tag{6}$$

которая, как легко видеть, удовлетворяет предположениям 3 из § 2.

2. Большой интерес представляет изучение частных случаев выполнения предположения 1, 2 из § 2 на функционалы J, J_δ, Ω и множество D. Предположим, что Z — рефлексивное банахово пространство. Следующие комплексы условий оказываются частными случаями предположений 1, 2.

Условия B. Топология τ — топология слабой сходимости в Z; D слабо замкнуто в Z; функционалы $\Omega(z)$, $J(z)$, $J_\delta(z)$ слабо полунепрерывны снизу на D (в частности, они могут быть слабо непрерывными); для $\Omega(z)$ выполнена при каждом $z \in D$ оценка $\Omega(z) \geqslant \bar{g}(\|z\|)$, где $\bar{g}(x)$ монотонно не убывает при $x \geqslant 0$ и $\bar{g}(+\infty) = +\infty$.

Условия C. Топология τ — топология слабой сходимости в Z; D выпукло и сильно замкнуто в Z; $\Omega(z)$, $J(z)$, $J_\delta(z)$ — выпуклые и сильно полунепрерывные снизу на D функционалы (в частности, возможен случай их сильной непрерывности); функционал $\Omega(z)$ удовлетворяет той же оценке, что и в условиях В.

Убедиться в том, что условия В, С обеспечивают выполнение предположения 1 из § 2, можно на основании определений и теорем, приведенных в § 0.3 (см. там теоремы 0.3.6—0.3.8). Установим, что из условий В, С следует также предположение 2 из § 2. Действительно, множество $\Omega_K=\{z\in D: \Omega(z)\leqslant K\}\neq\varnothing$ состоит из элементов, для которых выполнено неравенство $\bar{g}(\|z\|)\leqslant\Omega(z)\leqslant K$. Тогда существует такая константа K_1, что для любого $z\in\Omega_K$ верна оценка $\|z\|\leqslant K_1$. Чтобы убедиться в этом, предположим, что это не так. Тогда существует такая последовательность $\{z_n\}\subset\Omega_K$, для которой $\|z_n\|\to+\infty$ при $n\to\infty$ и для всякого n $\bar{g}(\|z_n\|)\leqslant K$. Но эти требования на z_n противоречивы вследствие того, что $\bar{g}(+\infty)=+\infty$. Из неравенства $\|z\|\leqslant K_1$, выполненного для элементов множества Ω_K, из условий на D и теорем 0.3.6, 0.3.7 ясно, что Ω_K представляет собой слабый компакт в Z.

Отметим, что условия С гарантируют применимость алгоритмов из § 7—9 к задачам выпуклого программирования.

Достаточные условия сильной и слабой полунепрерывности функционалов изучались в ряде работ, из которых упомянем [26, 32].

Выясним, какую сходимость приближений обеспечивают алгоритмы обобщенных принципов из § 7—9 при выполнении условий В или С. Из теорем 7.5, 8.4, 9.6 непосредственно получим, что для любой последовательности $\{\delta_n\}$, сходящейся к нулю при $n\to\infty$, соответствующая последовательность $\{z_n\}$ приближенных решений задачи (2.4) удовлетворяет предельным соотношениям

$$z_n \rightharpoonup \bar{Z},\quad \Omega(z_n)\to\bar{\Omega},\quad n\to\infty.$$

Этот результат можно уточнить, если рассмотреть задачу о поиске нормальных решений, т. е. предположить, что $\Omega(z)=g(\|z\|)$ и $g(x)$ удовлетворяет условию А. Тогда при $n\to\infty$

$$\|z_n\|=g^{-1}[\Omega(z_n)]\to g^{-1}[\bar{\Omega}]=g^{-1}[g(\|\bar{z}\|)]=\|\bar{z}\|.$$

Здесь $\bar{z}$ — произвольное Ω-оптимальное решение задачи (2.2).

Если теперь рефлексивное банахово пространство Z строго нормировано и обладает H-свойством, то, как указано в § 0.3, слабая сходимость $z_n \rightharpoonup \bar{Z}$ и сходимость норм $\|z_n\|\to\|\bar{z}\|$ при $n\to\infty$ обеспечивают сильную сходимость приближений $z_n\to\bar{Z}$. В частном случае, когда задача (2.2) имеет единственное нормальное решение $\bar{z}$, получим, что алгоритмы о.п.н., о.п.к., о.п.с.ф. дают сильную сходимость $z_n\to\bar{z}$ при $n\to\infty$. Аналогичный результат получается для нахождения обобщенного нормального решения относительно элемента $z_0\in D$, т. е. если $\Omega(z)=g(\|z-z_0\|)$.

Рассмотрим теперь более подробно случай, когда выполнены условия С.

Лемма 1. *Предположим, что в дополнение к условиям С функционал $\Omega(z)$ строго равномерно выпуклый на множестве D, а вспомогательная функция $f(x)$, кроме свойств, перечисленных в § 4, еще и выпуклая. Тогда:*

1) *множество* Z^* *решений задачи* (2.2), *предполагаемое непустым, будет выпуклым, а* Ω*-оптимальное решение* $\bar{z}$ *этой задачи будет единственным*;

2) *сглаживающий функционал* (4.1) *будет строго равномерно выпуклым на* D *при каждом* $\alpha>0$;

3) *множество* Z^α *решений задачи* (4.2) *при каждом* α *будет состоять из одного элемента.*

Доказательство. Выпуклость множества Z^* решений задачи минимизации выпуклого функционала J_δ на выпуклом множестве D следует из [32, с. 53, теорема 6]. Тогда задача (2.4) для нахождения Ω-оптимальных решений представляет собой задачу минимизации строго равномерно выпуклого функционала Ω на выпуклом множестве Z^* и потому имеет единственное решение. Установим второе утверждение леммы. Для этого возьмем произвольные элементы z_1, $z_2 \in D$. Тогда по определению строго равномерно выпуклого функционала Ω (см. § 0.3), и по определению выпуклого функционала J_δ имеем

$$\Omega\left(\frac{z_1+z_2}{2}\right) \leqslant \frac{\Omega(z_1)+\Omega(z_2)}{2} - \frac{1}{2} v(\|z_1-z_2\|),$$

$$J_\delta\left(\frac{z_1+z_2}{2}\right) \leqslant \frac{J_\delta(z_1)+J_\delta(z_2)}{2}.$$

Учитывая выпуклость вспомогательной функции f, получаем из этих неравенств

$$M^\alpha\left[\frac{z_1+z_2}{2}\right] = \alpha\Omega\left(\frac{z_1+z_2}{2}\right) + f\left[J_\delta\left(\frac{z_1+z_2}{2}\right)\right] \leqslant$$

$$\leqslant \frac{\alpha\Omega(z_1)+\alpha\Omega(z_2)}{2} + f\left[\frac{J_\delta(z_1)+J_\delta(z_2)}{2}\right] - \frac{\alpha}{2} v(\|z_1-z_2\|) \leqslant$$

$$\leqslant \frac{\alpha\Omega(z_1)+\alpha\Omega(z_2)}{2} + \frac{f[J_\delta(z_1)]+f[J_\delta(z_2)]}{2} - \frac{\alpha}{2} v(\|z_1-z_2\|) \leqslant$$

$$\leqslant \frac{M^\alpha(z_1)+M^\alpha(z_2)}{2} - \frac{\alpha}{2} v(\|z_1-z_2\|).$$

Найденное таким образом неравенство для $M^\alpha(z)$ и доказывает его строгую равномерную выпуклость. Единственность решения задачи (4.2) минимизации строго равномерно выпуклого сглаживающего функционала на выпуклом множестве D следует из теоремы 0.3.9.

Из леммы 1 с учетом следствия 6.1 получается непрерывность вспомогательных функций (6.1), а также функций (7.1), (7.2), (9.1).

Таким образом, при выполнении условий леммы 1 параметр регуляризации α_δ, выбранный по алгоритмам о.п.н., о.п.к. или о.п.с.ф., представляет собой обычное решение соответствующего уравнения (7.3), (8.1) или (9.2). Поэтому, как отмечалось в § 7—9, в этом случае нет необходимости в использовании правил отбора 7.1, 8.1 или 9.1: в качестве приближения к единственному Ω-оптимальному решению $\bar{z}$ берется единственный элемент z^{α_δ}, отвечающий выбранному значению α_δ. Как было указано, в случае, когда Z — рефлексивное, строго нормированное банахово пространство с H-свойством, алгоритмы гарантируют сильную сходимость приближений $z^{\alpha_{\delta_n}} \to \bar{z}$ при $\delta_n \to 0$.

В теоремах 7.4, 8.2 указаны условия, обеспечивающие единственность выбора параметра регуляризации по обобщенным принципам невязки и квазирешений. Наряду с требованием единственности решения задачи (4.2) при каждом $\alpha > 0$ эти условия содержат еще и неравенство $z^{\alpha_1} \neq z^{\alpha_2}$, которое должно выполняться для всех $\alpha_1 \neq \alpha_2$ ($\alpha_{1,2} > 0$). В случае дифференцируемости функционалов Ω, I_δ это условие можно сформулировать по-другому.

Лемма 2. *Предположим, что функционалы $\Omega(z)$, $I_\delta(z)$ непрерывно дифференцируемы на выпуклом множестве D банахова пространства Z и, кроме того, функционал $\Omega(z)$ выпуклый. Тогда равенство $z^{\alpha_1} = z^{\alpha_2}$ для различных α_1, $\alpha_2 > 0$ возможно лишь при $\alpha_1, \alpha_2 \geqslant \alpha^* \equiv \equiv \inf\{\alpha\colon \rho(\alpha) = \rho_\infty\} = \inf\{\alpha\colon \gamma(\alpha) = \Omega^*\}$. В этом случае $z^{\alpha_1} = z^{\alpha_2} \in Z_0$.*

Доказательство. Запишем необходимое условие минимума непрерывно дифференцируемого функционала $M^\alpha[z]$ на выпуклом множестве D (см. теорему 0.5.3):

$$\langle \{M^\alpha[z^\alpha]\}', z - z^\alpha \rangle = \langle \alpha\Omega'(z^\alpha) + I'_\delta(z^\alpha), z - z^\alpha \rangle \geqslant 0 \quad \forall z \in D.$$

Пусть теперь $\alpha_2 > \alpha_1 > 0$, $z^{\alpha_2} = z^{\alpha_1}$. Тогда из этого условия минимума получим

$$(\alpha_2 - \alpha_1) < \Omega'(z^{\alpha_2}), z - z^{\alpha_2} > \geqslant 0 \quad \forall z \in D.$$

Это неравенство для выпуклого $\Omega(z)$ представляет собой необходимое и достаточное условие минимума этого функционала на множестве D (см. теорему 0.4.3). Поэтому $z^{\alpha_1} = z^{\alpha_2} \in Z_0 = \{z \in D\colon \Omega(z) = \Omega^*\}$ и, значит, $\gamma(\alpha_1) = \gamma(\alpha_2) = \Omega^*$. Но в силу монотонности функции $\gamma(\alpha)$ это равенство возможно лишь при α_1, $\alpha_2 \geqslant \alpha^*$, если α^* конечно, или невозможно вообще. Лемма доказана.

Из леммы 2 и теорем 7.4, 8.2 легко получается

Теорема 1. *Пусть D — выпуклое, сильно замкнутое множество в рефлексивном банаховом пространстве Z. Предположим, что функционалы $J_\delta(z)$, $\Omega(z)$ непрерывно дифференцируемы на D, причем $J_\delta(z)$ — выпуклый, а $\Omega(z)$ — строго равномерно выпуклый на D функционал. Считаем также, что функция $f(x) \in \mathcal{F}^m$ ($m \geqslant 1$) выпуклая. Если выполнены неравенства*

$$\rho_0 = f(J^*_\delta) - f[\Psi(\delta, \Omega_\delta) + \lambda_\delta] < 0,$$
$$\rho_\infty = \nu_\delta - f[\Psi(\delta, \Omega^*) + \lambda_\delta] > 0, \quad \Omega^* < \omega_\delta < \Omega_\delta,$$

то уравнения (7.3), (8.1), *используемые для выбора параметра регуляризации в алгоритмах о.п.н., о.п.к., имеют единственное обычное решение* $\alpha_\delta > 0$.

Действительно, существование обычных решений $\alpha_\delta > 0$ уравнений (7.3), (8.1) доказывается, как в теоремах 7.4, 8.2, исходя из результата леммы 1 о единственности экстремали сглаживающего функционала. При этом числа α_δ обязаны удовлетворять неравенству $\alpha_\delta < \alpha^*$, так как иначе выполнялись бы равенства $\rho(\alpha_\delta) = 0 = \rho_\infty$, $\gamma(\alpha_\delta) = \omega_\delta = \Omega^*$, противоречащие условию теоремы. Следовательно, по лемме 2 в случае существования нескольких решений уравнения (7.3) или (8.1) им должны отвечать различные экстремали задачи (4.2). Это в свою очередь позволяет развить схему доказательства единственности решений α_δ, примененную в теоремах 7.4, 8.2.

3. Приведем некоторые конкретные примеры стабилизирующих функционалов $\Omega(z)$, которые в зависимости от выбора пространства Z и топологии τ порождают различные типы сходимости приближений в алгоритмах из § 7—9.

Пример 2. Предположим, что $Z=D=H^\lambda\,[a, b]$ — пространство функций, удовлетворяющих на отрезке $[a, b]$ условию Гёльдера с показателем λ $(0<\lambda\leqslant 1)$ (см. § 0.4). Положим

$$\Omega_1(z)=\|z\|_{H^\lambda}=$$
$$=\|z\|_{C[a,b]}+\sup\{|z(x)-z(y)|/|x-y|^\lambda\colon x, y\in[a, b], x\neq y\}. \tag{7}$$

Зададим в Z топологию равномерной сходимости. Тогда функционал $\Omega_1(z)$ будет τ-секвенциально полунепрерывным снизу в Z. Действительно, если последовательность $\{z_n\}\subset H^\lambda[a, b]\subset C[a, b]$ равномерно сходится к элементу $z_0\in H^\lambda$, то для любого $x, y\in[a, b]$ $(x\neq y)$

$$\lim_{n\to\infty}\Omega_1(z_n)\geqslant\lim_{n\to\infty}\{\|z_n\|_C+|z_n(x)-z_n(y)|/|x-y|^\lambda\}=$$
$$=\|z_0\|_C+|z_0(x)-z_0(y)|/|x-y|^\lambda.$$

Отсюда в силу произвольности x, y получается доказываемое неравенство

$$\varliminf_{n\to\infty}\Omega_1(z_n)\geqslant\|z_0\|_C+\sup_{\substack{x,\,y\in[a,\,b]\\ x\neq y}}\{|z_0(x)-z_0(y)|/|x-y|^\lambda\}\equiv\Omega_1(z_0).$$

Установим теперь τ-секвенциальную компактность множества $\Omega_C=\{z\in H^\lambda\colon \|z\|_{H^\lambda}\leqslant C\}$ $(C=\mathrm{const}>0)$. Из неравенства

$$\|z\|_C+|z(x)-z(y)|/|x-y|^\lambda\leqslant\|z\|_{H^\lambda}\leqslant C,$$

выполненного для любого элемента $z\in\Omega_C$ и всяких $x, y\in[a, b]$ $(x\neq y)$, следует равномерная ограниченность множества Ω_C в пространстве $C\,[a, b]$.

Из этого же неравенства получается оценка $|z(x)-z(y)|\leqslant C|x-y|^\lambda$, означающая равностепенную непрерывность функций из множества Ω_C (см. § 0.4). Тогда τ-компактность множества Ω_C следует из теоремы Арцела.

Использование в алгоритмах § 7—9 приведенных в этом примере объектов (Z, τ), Ω_1 даст в силу теорем 7.5, 8.4, 9.6, равномерную сходимость приближений.

Пример 3. Зададим $Z=D=L_1\,[a, b]$ и будем считать, что τ — топология сильной сходимости в Z. Определим функционал

$$\Omega_2(z)=\|z\|_{L_1[a,b]}+\sup\left\{\frac{1}{h}\int_a^b|z(x+h)-z(x)|\,dx\colon 0<h\leqslant h_0\right\},$$

полагая $z(x)=0$, если $x\notin[a, b]$. Этот функционал будет τ-секвенциально полунепрерывен снизу: для любой последовательности

$\{z_n\}\subset L_1[a,b]$, сходящейся при $n\to\infty$ по норме к элементу $z_0\in L_1[a,b]$, выполнено неравенство

$$\varliminf_{n\to\infty}\Omega_2(z_n)\geqslant\varliminf_{n\to\infty}\frac{1}{h}\int\limits_a^b|z_n(x+h)-z_n(x)|\,dx+$$
$$+\varliminf_{n\to\infty}\|z_n\|_{L_1}=\frac{1}{h}\int\limits_a^b|z_0(x+h)-z_0(x)|\,dx+\|z_0\|_{L_1},\quad 0<h\leqslant h_0.$$

Отсюда в силу произвольности числа h и получается искомое неравенство

$$\varliminf_{n\to\infty}\Omega_2(z_n)\geqslant\|z_0\|_{L_1}+\sup\left\{\frac{1}{h}\int\limits_a^b|z_0(x+h)-z_0(x)|\,dx\colon 0<h\leqslant h_0\right\}=\Omega_2(z_0).$$

Докажем τ-секвенциальную компактность множества $\Omega_C=\{z\in\in L_1[a,b]\colon\Omega_2(z)\leqslant C\}$. Из неравенства

$$\|z\|_{L_1}+\frac{1}{h}\int\limits_a^b|z(x+h)-z(x)|\,dx\leqslant\Omega_2(z)\leqslant C$$

следует, что для любого $z\in\Omega_C$

$$\|z\|_{L_1}\leqslant C,\quad\int\limits_a^b|z(x+h)-z(x)|\,dx\leqslant Ch,\quad 0<h\leqslant h_0.$$

Из этих условий по критерию М. Рисса (см. теорему 0.4.5) и получается искомая компактность.

Пример 4. Пусть снова $Z=L_1[a,b]$, D — множество функций из $L_1[a,b]$, равномерно ограниченных константой M на отрезке $[a,b]$: $|z(x)|\leqslant M$ $(x\in[a,b])$. Будем считать, что τ — топология сходимости почти всюду на $[a,b]$. Определим

$$\Omega_3(z)=\sup\left\{\frac{1}{h}\int\limits_a^b|z(x+h)-z(x)|\,dx\colon 0<h\leqslant h_0\right\}.$$

Функционал $\Omega_3(z)$ τ-секвенциально полунепрерывен снизу на множестве D, что непосредственно следует из следствия теоремы Лебега (см. теорему 0.4.8). Убедимся, что Ω_C — τ-секвенциально компактное множество. Как ясно из примера 3, множество Ω_C компактно в смысле сходимости в L_1 $[a,b]$. Тогда (см. § 0.4) оно компактно в смысле сходимости по мере и почти всюду.

Использование функционала Ω_3 и пространства (Z,τ) из примера 4 в алгоритмах, изученных в § 7—9, обеспечивает сходимость приближений почти всюду.

Пример 5. Возьмем $Z=D=V[a,b]$ — пространство функций с ограниченной вариацией на отрезке $[a,b]$ (см. § 0.4 и гл. 4).

В качестве τ возьмем топологию поточечной сходимости ограниченных в $V[a, b]$ последовательностей. По лемме 4.1.2′ функционал

$$\Omega_4(z)=|z(a)|+\bigvee_a^b(z)\equiv\|z\|_V$$

τ-секвенциально полунепрерывен снизу на $V[a, b]$. Множество Ω_C, для элементов которого выполнена оценка $\|z\|_V\leqslant C$, будет τ-секвенциально компактным, т. е. компактным в смысле поточечной сходимости, по теореме Хелли (см. теорему 0.4.15). Предположим, что задача (2.2) рассматривается над пространством (Z, τ) из данного примера и имеет единственное нормальное в $V[a, b]$ решение $\bar{z}$. Тогда алгоритмы о.п.н., о.п.к., о.п.с.ф., использующие функционал Ω_4, обеспечивают поточечную сходимость на $[a, b]$ приближений $z^{\alpha_\delta}(x)$ к $\bar{z}(x)$ при $\|\delta\|\to 0$, а также сходимость вариаций

$$\bigvee_a^b(z^{\alpha_\delta})\to\bigvee_a^b(\bar{z}),\quad \|\delta\|\to 0.$$

Если, кроме того, $\bar{z}\in C[a, b]$, то, как будет показано в следствии 2 из теоремы 4.1.1, сходимость приближений к $\bar{z}(x)$ будет равномерной на отрезке $[a, b]$.

Имеются и другие примеры используемых в практических вычислениях функционалов Ω (см. [108, 182]).

§ 12. Связь обобщенных принципов невязки и квазирешений с обобщенным методом невязки и методом квазирешений

Сформулируем задачу определения Ω-оптимальных решений (2.4) в несколько иной форме:

$$\operatorname{Arg\,inf}\{\Omega(z)\colon\ z\in D,\ \ J(z)=J^*\}=\bar{Z}. \tag{1}$$

До сих пор мы рассматривали схему приближенного решения такой задачи, изложенную в § 4 и базирующуюся на минимизации сглаживающего функционала. Возьмем, однако, другой вариационный подход к решению (1), использующий приближенные данные $\{J_\delta, \Psi, \delta\}$. Он основан на рассмотрении экстремальной задачи, «близкой» к (1): найти элементы $z_\delta=z_\delta(J_\delta, \Psi, \delta)$, для которых

$$\Omega(z_\delta)=\inf\{\Omega(z)\colon z\in D,\ J_\delta(z)\leqslant\Psi(\delta, \Omega(z))+\lambda_\delta\}\equiv\tilde{\Omega}_\delta. \tag{2}$$

Если эта задача имеет решения, то любое из них принимается в качестве приближения к множеству $\bar{Z}$. Такой метод построения приближенных решений задачи (1) называется *обобщенным методом невязки* (о.м.н.) для экстремальных задач.

Легко убедиться, что даже при выполнении предположений 1—3 из § 2 задача (2) может оказаться не разрешимой. Рассмотрим простой пример.

Пример 1. Пусть $Z=\mathbf{R}$, а топология τ в Z порождается нормой $\|z\|=|z|$ $(z\in\mathbf{R})$. Зададим $D=\{z\in\mathbf{R}\colon z\geqslant 0\}$ и определим на D функционалы

$$J(z)=\{1\colon z\neq 1;\ 0\colon z=1\},\qquad \Omega(z)=\{0\colon z=0;\ 1+z\colon z>0\}.$$

Выделим класс $\mathscr{F}$ допустимых приближенных функционалов $J_\delta(z)$ с помощью условия аппроксимации $|J_\delta(z)-J(z)|\leqslant\Psi(\delta,\Omega(z))=\delta\Omega(z)$, где положительное число δ определяет погрешность аппроксимации. Предположим теперь, что в нашем распоряжении имеются приближенные данные для решения задачи (1): $\{J_\delta,\Psi,\delta\}=\{J,\delta\Omega,\delta\}$, т. е. оказалось, что приближенный функционал совпадает с точным. Рассмотрим задачу (2), отвечающую этим данным для числа δ, подчиненного неравенству $1/3<\delta<1/2$. Тогда

$$\lambda_\delta=\inf\{J_\delta(z)+\Psi(\delta,\Omega(z)):\ z\in D\}=\min(1,\ 1+\delta,\ 2\delta)=2\delta.$$

Поэтому ограничение задачи (2) имеет вид

$$J(z)\leqslant 2\delta+\delta\Omega(z)=\{2\delta:\ z=0;\ 3\delta+\delta z:\ z>0\}=\pi_\delta(z).$$

Отсюда ясно, что множество элементов z, удовлетворяющих этому ограничению, есть интервал $(0,+\infty)$ и функционал Ω не может достигать на нем своей точкой нижней грани. Иллюстрацией к примеру 1 служит рис. 4.

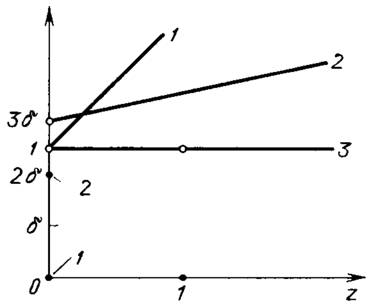

Рис. 4. Графики функционалов: *1* — $\Omega(z)$; *2* — $\pi_\delta(z)$; *3* — $J_\delta(z)$

Теорема 1. *Если выполнены предположения* 1—3 *из* § 2 *и в дополнение к ним функционал* Ω *τ-секвенциально непрерывен на* D, *то задача* (2) *имеет решение.*

Доказательство. Определим для удобства множество ограничений задачи (2): $Z(\delta)=\{z\in D:\ J_\delta(z)\leqslant\Psi(\delta,\Omega(z))+\lambda_\delta\}$. Это множество непусто, так как

$$J_\delta(\bar z)\leqslant J(\bar z)+\Psi(\delta,\Omega(\bar z))\leqslant \leqslant\lambda_\delta+\Psi(\delta,\Omega(\bar z))\quad\forall\bar z\in\bar Z$$

(см. теорему 3.1). Введем далее множество $\bar Z(\delta)=\{z\in Z(\delta):\ \Omega(z)\leqslant\bar\Omega\}$, также содержащее элементы $\bar z$. Тогда очевидно, что

$$\inf\{\Omega(z):\ z\in Z(\delta)\}=\inf\{\Omega(z):\ z\in\bar Z(\delta)\}. \tag{3}$$

Множество $\bar Z(\delta)$ является τ-секвенциально компактным. Действительно, любая последовательность $\{z_n\}\subset\bar Z(\delta)$ удовлетворяет неравенству $\Omega(z_n)\leqslant\bar\Omega$, и вследствие предположения 2 из § 2 из нее можно извлечь подпоследовательность $\{z_{n_k}\}$, которая τ-сходится при $k\to\infty$ к элементу $z_0\in D$. В силу τ-секвенциальной полунепрерывности функционала J_δ и непрерывности Ω и Ψ получим

$$\Omega(z_0)=\lim_{k\to\infty}\Omega(z_{n_k})\leqslant\bar\Omega,$$

$$J_\delta(z_0)\leqslant\varliminf_{k\to\infty}J_\delta(z_{n_k})\leqslant\varliminf_{k\to\infty}[\lambda_\delta+\Psi(\delta,\Omega(z_{n_k}))]=\lambda_\delta+\Psi(\delta,\Omega(z_0)).$$

Таким образом, $z_0 \in \bar{Z}(\delta)$ и $z_{n_k} \xrightarrow{\tau} z_0 \in \bar{Z}(\delta)$. Это доказывает τ-секвенциальную компактность множества $\bar{Z}(\delta)$. Такая компактность множества $\bar{Z}(\delta)$ вместе с τ-секвенциальной непрерывностью функционала Ω на этом множестве по теореме типа Вейерштрасса (см. теорему 0.1.2) дает существование решения $\tilde{z}_\delta \in \bar{Z}(\delta)$ экстремальной задачи

$$\operatorname{Arg}\inf\{\Omega(z)\colon z \in \bar{Z}(\delta)\}.$$

Тогда в силу (3) будет разрешима и задача (2), а элемент $\tilde{z}_\delta$ будет ее решением, так как $\tilde{z}_\delta \in \bar{Z}(\delta) \subset Z(\delta)$.

Условие τ-секвенциальной непрерывности функционала Ω не является необходимым для разрешимости задачи (2). В ряде случаев эта задача имеет решение и для τ-секвенциально полунепрерывных снизу на D функционалов Ω.

Теорема 2. *Пусть выполнены предположения* 1—3 *из* § 2 *и уравнение* (7.3) *для нахождения параметра регуляризации по обобщенному принципу невязки имеет обычное решение* $\alpha_\delta > 0$. *Тогда задача* (2) *имеет решение.*

Доказательство. Существование обычного решения $\alpha_\delta > 0$ уравнения (7.3) означает по определению обычного решения и по определению функции $\rho(\alpha)$, входящей в (7.3), что существует такая экстремаль z^{α_δ} задачи (4.2) для $\alpha = \alpha_\delta$, для которой выполнено равенство

$$\rho(\alpha_\delta) = f\left[J_\delta(z^{\alpha_\delta})\right] - f\left[\lambda_\delta + \Psi(\delta, \Omega(z^{\alpha_\delta}))\right] = 0. \tag{4}$$

Из этого равенства и возрастания функции f ясно, что $z^{\alpha_\delta} \in Z(\delta)$.

Покажем теперь, что этот элемент z^{α_δ} есть решение задачи (2) Для этого возьмем произвольный элемент $\tilde{z} \in Z(\delta)$. Тогда для него

$$I_\delta(\tilde{z}) \equiv f\left[J_\delta(\tilde{z})\right] \leqslant f\left[\lambda_\delta + \Psi(\delta, \Omega(\tilde{z}))\right]. \tag{5}$$

Учитывая (4), (5), а также экстремальность элемента z^{α_δ} как решения задачи (4.2), получаем

$$\begin{aligned} \alpha_\delta\Omega(z^{\alpha_\delta}) + f\left[\lambda_\delta + \Psi(\delta, \Omega(z^{\alpha_\delta}))\right] &= \alpha_\delta\Omega(z^{\alpha_\delta}) + I_\delta(z^{\alpha_\delta}) = \\ &= M^{\alpha_\delta}\left[z^{\alpha_\delta}\right] \leqslant M^{\alpha_\delta}\left[\tilde{z}\right] = \alpha_\delta\Omega(\tilde{z}) + I_\delta(\tilde{z}) \leqslant \\ &\leqslant \alpha_\delta\Omega(\tilde{z}) + f\left[\lambda_\delta + \Psi(\delta, \Omega(\tilde{z}))\right], \end{aligned} \tag{6}$$

т. е.

$$\Phi\left[\Omega(z^{\alpha_\delta})\right] \leqslant \Phi\left[\Omega(\tilde{z})\right], \tag{7}$$

где $\Phi(\Omega) \equiv \alpha_\delta\Omega + f\,[\lambda_\delta + \Psi(\delta, \Omega)]$. Из неубывания функции Ψ по второму аргументу при фиксированном δ, возрастания функции f и неравенства $\alpha_\delta > 0$ следует, что функция $\Phi(\Omega)$ — монотонно возрастающая при $\Omega \geqslant \Omega^*$. Поэтому неравенство (7) приводит к оценке $\Omega(z^{\alpha_\delta}) \leqslant \Omega(\tilde{z})$ для произвольного $\tilde{z}$ из множества $Z(\delta)$. Это и означает, что z^{α_δ} — решение задачи (2).

Теорема 2 показывает, что при выполнении ее условий все приближенные решения задачи (1), получаемые по алгоритму о.п.н.

для экстремальных задач, удовлетворяют также условиям обобщенного метода невязки. В некоторых случаях о.п.н. и о.м.н. оказываются эквивалентными.

Теорема 3. *Предположим, что выполнены условия теоремы* 7.4. *Тогда элемент* z^{α_δ}, *выбираемый по алгоритму о.п.н., является единственным решением задачи* (2).

Доказательство. В условиях теоремы 7.4 уравнение (7.3) имеет единственное решение $\alpha_\delta>0$, и этому значению параметра регуляризации соответствует единственная экстремаль задачи (4.2) z^{α_δ}. Поэтому для любого элемента $\tilde{z}\in Z(\delta)$, отличного от z^{α_δ}, будет выполнено неравенство

$$\alpha_\delta\Omega\left(z^{\alpha_\delta}\right)+I_\delta\left(z^{\alpha_\delta}\right)=M^{\alpha_\delta}\left[z^{\alpha_\delta}\right]<M^{\alpha_\delta}\left[\tilde{z}\right]=\alpha_\delta\Omega\left(\tilde{z}\right)+I_\delta\left(\tilde{z}\right).$$

Это неравенство представляет собой усиление оценки (6). Поэтому из него с использованием соотношений (4), (5), как при доказательстве теоремы 2, можно получить усиленное неравенство (7): $\Phi\ [\Omega(z^{\alpha_\delta})]<\Phi\ [\Omega(\tilde{z})]$, откуда вследствие возрастания функции $\Phi(\Omega)$ следует оценка $\Omega(z^{\alpha_\delta})<\Omega(\tilde{z})$. Из нее в силу произвольности элемента $\tilde{z}\in Z(\delta)$ и следует, что z^{α_δ} — единственное решение задачи (2).

Таким образом, алгоритмы о.п.н. и о.м.н. эквивалентны при выполнении условий теоремы 7.4, и приближенное решение, получаемое на их основе, единственно. Если условия теоремы 7.4 нарушаются, то вопрос об эквивалентности этих алгоритмов решается в общем случае отрицательно. Соответствующие примеры приведены в § 2.5.

Изучим вопрос о сходимости приближенных решений задачи (1), полученных по о.м.н.

Теорема 4. *Пусть для произвольной последовательности* $\{\varepsilon_n\}$ ($\varepsilon_n>0$), *сходящейся к нулю при* $n\to\infty$, *и для любой последовательности* $\{\delta_n\}$ *такой, что* $\|\delta_n\|\to 0$ *при* $n\to\infty$, *элементы* $z_n\in Z(\delta_n)$ *находятся как* ε-*приближенные решения задачи* (2): $\tilde{\Omega}_{\delta_n}\leqslant\Omega(z_n)\leqslant\tilde{\Omega}_{\delta_n}+\varepsilon_n$. *Тогда при* $n\to\infty$ *выполнены предельные соотношения* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$.

Доказательство. При определении множества $Z(\delta)$ в теореме 1 было отмечено, что $\bar{Z}\subset Z(\delta_n)$ для всякого δ_n. Поэтому $\tilde{\Omega}_{\delta_n}\leqslant\Omega(\bar{z})=\bar{\Omega}$ для любого $\bar{z}\in\bar{Z}$, и в силу условия теоремы 4 $\Omega(z_n)\leqslant\bar{\Omega}+\varepsilon_n$. Из этого неравенства ясно, что

$$\overline{\lim_{n\to\infty}}\ \Omega(z_n)\leqslant\bar{\Omega}. \tag{8}$$

Поскольку $z_n\in Z(\delta_n)$, то из условия аппроксимации (2.3) и монотонности по второму аргументу функции Ψ следует

$$J^*-\Psi(\delta_n,\ \bar{\Omega}+\varepsilon_n)=J(\bar{z})-\Psi(\delta_n,\ \Omega(\bar{z})+\varepsilon_n)\leqslant J(z_n)-\Psi(\delta_n,\ \Omega(z_n))\leqslant$$
$$\leqslant J_{\delta_n}(z_n)\leqslant\lambda_{\delta_n}+\Psi(\delta_n,\ \Omega(z_n))\leqslant\lambda_{\delta_n}+\Psi(\delta_n,\ \bar{\Omega}+\varepsilon_n).$$

Перейдем к пределу при $n\to\infty$ в этом неравенстве. При этом учтем, что по теореме 3.1 имеем $\lambda_{\delta_n}\to J^*$, а также используем сходимость $\Psi(\delta_n, \bar{\Omega}+\varepsilon_n)\to 0$, вытекающую из свойств функции Ψ и из оценки $0\leqslant\Psi(\delta_n, \bar{\Omega}+\varepsilon_n)\leqslant\Psi(\delta_n, \Omega_{\max})$, $\Omega_{\max}\equiv\bar{\Omega}+\max\varepsilon_n$. Тогда получим

$$\lim_{n\to\infty} J_{\delta_n}(z_n)=J^*. \tag{9}$$

Соотношения (8), (9) означают, что выполнены условия леммы 5.1, по которой и оказываются справедливыми доказываемые в теореме 4 сходимости.

Отметим, что теорема 4 не требует разрешимости задачи (2), хотя и включает в себя эту возможность.

Обратимся теперь к изучению связи алгоритма о.п.к. и так называемого *метода квазирешений* (м.к.) для экстремальных задач. Он заключается в следующем. В качестве приближенного решения задачи (1) принимается любое решение z_δ экстремальной задачи

$$J_\delta(z_\delta)=\inf\{J_\delta(z)\colon z\in D,\ \Omega(z)\leqslant\omega_\delta\}\equiv J^*(\delta). \tag{10}$$

Теорема 5. *Если выполнены предположения* 1—3 *из* § 2 *и число* ω_δ *удовлетворяет неравенству* $\omega_\delta\geqslant\Omega^*$, *то задача* (10) *разрешима.*

Доказательство. Достаточно заметить, что множество

$$\Omega_{\omega_\delta}\equiv\{z\in D\colon \Omega(z)\leqslant\omega_\delta\}$$

непусто при выполнении неравенства $\omega_\delta\geqslant\Omega^*$ (см. лемму 5.12). Поэтому из предположения 2 из § 2 вытекает его τ-секвенциальная компактность. Это вместе с τ-секвенциальной полунепрерывностью снизу функционала J_δ по теореме 0.1.2 обосновывает разрешимость задачи (10).

Теорема 6. *Пусть выполнены предположения* 1—3 *из* § 2 *и число* ω_δ *взято так, что* $\omega_\delta\geqslant\bar{\Omega}$ *при каждом* δ, *причем* $\omega_\delta\to\bar{\Omega}$ *при* $\Delta\to 0$. *Пусть далее для произвольной последовательности* $\{\varepsilon_n\}$ $(\varepsilon_n>0)$ *такой, что* $\varepsilon_n\to 0$ *при* $n\to\infty$, *и для любой последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n\to\infty$, *элементы* z_n *находятся как* ε-*приближенные решения задачи* (10), *т. е. удовлетворяют неравенствам* $J^*(\delta_n)\leqslant$ $\leqslant J_{\delta_n}(z_n)\leqslant J^*(\delta_n)+\varepsilon_n$, $\Omega(z_n)\leqslant\omega_{\delta_n}$. *Тогда* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Доказательство. Покажем сначала, что $J^*(\delta_n)\to J^*$ при $n\to\infty$. Для этого возьмем последовательность $\{z_{\delta_n}\}$ решений задач типа (10) для указанных чисел δ_n. Из условия (2.3) и неубывания функции Ψ по аргументу Ω получим

$$J^*(\delta_n)=J_{\delta_n}(z_{\delta_n})\geqslant J(z_{\delta_n})-\Psi(\delta_n, \Omega(z_{\delta_n}))\geqslant J^*-\Psi(\delta_n, \omega_{\delta_n}).$$

По условиям теоремы $\Omega(\bar{z})=\bar{\Omega}\leqslant\omega_\delta$ для любого $\bar{z}\in\bar{Z}$, т. е. $\bar{Z}\subset\Omega_{\omega_\delta}$. Поэтому из (10) следует также

$$J_{\delta_n}(z_{\delta_n})=J^*(\delta_n)\leqslant J_{\delta_n}(\bar{z})\leqslant J(\bar{z})+\Psi(\delta_n, \Omega(\bar{z}))=J^*+\Psi(\delta_n, \bar{\Omega}).$$

Переходя в полученных неравенствах к пределу при $n\to\infty$ и учитывая сходимости $\Psi(\delta_n, \omega_{\delta_n})\to 0$, $\Psi(\delta_n, \bar{\Omega})\to 0$, приходим к выводу о том, что $J^*(\delta_n)\to J^*$ при $n\to\infty$. Тогда из условий на

ε-приближенные решения z_n получаются соотношения (8), (9). Они в свою очередь по лемме 5.1 и дают утверждение теоремы 6.

В частности, теорема 6 обосновывает, что метод квазирешений для экстремальных задач, основанный на использовании точных решений задачи (10), является регуляризующим алгоритмом.

Возникает вопрос о связи метода квазирешения и обобщенного принципа квазирешений для экстремальных задач.

Теорема 7. *Пусть выполнены условия теоремы* 8.2. *Тогда приближенное решение* z^{α_δ} *задачи* (1), *определяемое по о.п.к., является единственным решением задачи* (10).

Доказательство. Из теоремы 8.2 следует, что существует единственное значение $\alpha_\delta > 0$ параметра регуляризации и единственное приближение z^{α_δ}, получаемые по алгоритму о.п.к. При этом условие (8.1) принимает вид

$$\gamma(\alpha_\delta) = \Omega(z^{\alpha_\delta}) = \omega_\delta. \tag{11}$$

Возьмем теперь произвольный элемент $\tilde{z}$, отличный от z^{α_δ} и удовлетворяющий неравенству $\Omega(\tilde{z}) \leqslant \omega_\delta$. Тогда вследствие единственности решения задачи (4.2) при $\alpha = \alpha_\delta$ получим из (11)

$$\alpha_\delta \omega_\delta + I_\delta(z^{\alpha_\delta}) = \alpha_\delta \Omega(z^{\alpha_\delta}) + I_\delta(z^{\alpha_\delta}) = M^{\alpha_\delta}[z^{\alpha_\delta}] < M^{\alpha_\delta}[\tilde{z}] =$$
$$= \alpha_\delta \Omega(\tilde{z}) + I_\delta(\tilde{z}) \leqslant \alpha_\delta \omega_\delta + I_\delta(\tilde{z}),$$

т. е. $I_\delta(z^{\alpha_\delta}) < I_\delta(\tilde{z})$. Таким образом, элемент z^{α_δ}, входя, согласно (11), во множество Ω_{ω_δ}, является единственным решением задачи (10).

Теорема 7 дает условия эквивалентности о.п.к. и м.к. для экстремальных задач. При нарушении условия единственности экстремали сглаживающего функционала при каждом $\alpha > 0$, которое фигурирует в теореме 8.2, вопрос об эквивалентности о.п.к. и м.к. решается, вообще говоря, отрицательно. Соответствующий пример приведен в § 2.5.

Замечания. Обобщенный метод невязки был предложен для решения нелинейных операторных уравнений в работе [60] и представляет собой развитие метода невязки, восходящего к работам [239, 86] (см. также [36, 40, 76, 77, 142, 152, 203]). В дальнейшем о.м.к. исследовался для линейных операторных уравнений в [88, 155, 183, 187], а для нелинейных операторных уравнений в [160, 162, 230] (случай метрических пространств) и в [105, 106, 123, 224, 226] (случай рефлексивных банаховых пространств). Для экстремальных задач о.м.н. изучался в [133] в общем случае, а в [142] в специальном случае равномерной аппроксимации точного функционала приближенным: $\Psi(\delta, \Omega) = \Psi(\delta)$. Связь алгоритмов о.п.н. и о.м.н. исследовалась для операторных уравнений в работах [88, 123, 162, 226] и др.

Метод квазирешений предложен для решения операторных уравнений в работах В. К. Иванова (см. [84—88]). Для экстремальных задач он изучался в статьях [24, 221] (случай выпуклого функционала), [8, 193] (выпуклые задачи в банаховых пространствах), [140, 142, 203] и др.

Вопрос о связи принципа квазирешений и метода квазирешения рассматривался для случая операторных уравнений в [86, 88, 187] и др.

§ 13. Оптимальность по порядку точности алгоритмов о.п.н., о.п.к., о.п.с.ф. для решения экстремальных задач

Как известно [42, 88], спецификой некорректно поставленных задач является невозможность получения априорных оценок точности регуляризирующих алгоритмов на достаточно общих множествах входных данных без дополнительных существенных предположений о структуре этих множеств. По этой причине оценки точности методов решения некорректных задач чаще всего получают с привлечением информации о принадлежности данных задачи и (или) ее решения некоторому компакту [42, 43, 53, 88, 109—112]. Использование такой информации делает в ряде случаев решаемую задачу корректной [189]. Получаемые таким образом оценки, как правило, не являются конструктивными: их нельзя, вообще говоря, использовать при оценке точности приближенных решений конкретных прикладных, некорректно поставленных задач. Вследствие этого для практических оценок точности приближений, получаемых с помощью регуляризирующих алгоритмов, используется обычно методика квазиреального эксперимента (модельных примеров) [82, 92, 204, 205]. Для этих целей в последнее время развита также теория апостериорных оценок погрешности некорректных задач [43, 52, 53].

Тем не менее оценки точности регуляризирующих алгоритмов на компакте (на слабом компакте) имеют большое теоретическое значение в плане установления оптимальности данного алгоритма по порядку точности. Теория исследования алгоритмов на оптимальность по порядку точности развивалась в основном для случая линейных некорректных задач в работах [88, 185, 187, 188].

В данном параграфе будут даны оценки точности приближенных решений и установлена оптимальность порядка этой точности для алгоритмов обобщенных принципов невязки, квазирешений и сглаживающего функционала, а также методов невязки и квазирешений применительно к решению экстремальных задач типа (2.2) [133].

Сформулируем постановку задачи. Пусть Z — банахово пространство, а функционал $J_0(z)$ определен на некотором множестве $D\subset Z$ $(D\neq\varnothing)$ и ограничен там снизу. Рассмотрим экстремальную задачу: найти элементы $\bar z\in D$, такие, что

$$J_0(\bar z)=\inf\{J_0(z)\colon z\in D\}\equiv J_0^*. \tag{1}$$

Будем считать, что эта задача имеет единственное решение $\bar z_0\in D$. Предположим далее, что вместо точного функционала J_0 известен приближенный функционал $J_\delta(z)$, определенный на D, для которого выполнено условие аппроксимации

$$|J_0(z)-J_\delta(z)|\leqslant\psi_0(\delta,\ \|z\|)\quad \forall z\in D. \tag{2}$$

Здесь по-прежнему конечномерный вектор $\delta\in\mathbf{R}^k$, евклидова норма $\Delta=\|\delta\|\neq0$ которого определяет степень близости J_0 и J_δ, считается известным. Известна также мера аппроксимации ψ_0 функционала J_0 функционалом J_δ; она удовлетворяет предположениям 3 из § 2, и в дополнение к ним функция $\psi_0(\delta, x)$ возрастает по аргументу $x\geqslant0$ при каждом δ $(\delta\neq\theta)$.

Будем рассматривать вопрос об устойчивом в Z приближенном нахождении решения задачи (1): по совокупности данных $\{J_\delta, \delta, \psi_0\}$ найти элемент $z_\delta \in D$ такой, что $z_\delta \to \bar{z}_0$ при $\Delta \to 0$.

Предположим, что множество D сильно замкнуто в Z, а функционалы J_0, J_δ полунепрерывны снизу на D. Кроме того, предположим, что о решении $\bar{z}_0$ имеется априорная информация следующего вида: элемент $\bar{z}_0 \in Z$ истокообразно представим с помощью оператора B, т. е. $\bar{z}_0 = B\bar{v}$. Здесь B — линейный, взаимно однозначный, вполне непрерывный оператор, действующий из рефлексивного банахова пространства V в Z, а $\bar{v} \in V$, $\bar{v} \neq 0$.

В качестве оператора B, в частности, можно взять оператор компактного вложения пространства V в Z. Такая постановка задачи носит название *схемы компактного вложения*.

Если теперь ввести множество $\mathcal{V} \equiv \{v \in V\colon Bv \in D\}$, содержащее элемент $\bar{v}$, то задачу (1) можно переформулировать так: найти элемент $\bar{v} \in \mathcal{V}$, для которого

$$J_0(B\bar{v}) = \inf\{J_0(Bv)\colon v \in \mathcal{V}\}. \tag{1'}$$

Тогда для приближенного решения задачи (1), (1′) достаточно найти такие элементы $v_\delta \in \mathcal{V}$, которые слабо сходятся при $\Delta \to 0$ к $\bar{v}$: $v_\delta \rightharpoonup \bar{v}$. В этом случае $z_\delta \equiv Bv_\delta \xrightarrow{z} B\bar{v} \equiv \bar{z}_0$.

Для решения такой задачи будут использованы алгоритмы обобщенных принципов невязки, квазирешений и сглаживающего функционала. Для этого введем регуляризирующий функционал $\Omega(v)$ в виде $\Omega(v) = g(\|v\|)$, где функция $g(x)$ считается определенной при $x \in [0, +\infty)$, неотрицательной, непрерывной, монотонно возрастающей и удовлетворяющей требованию $g(+\infty) = +\infty$. Для удобства вместо приближенного функционала $J_\delta(Bv)$ будем использовать новый приближенный функционал $J_\delta(Bv) + \Psi(\delta, \Omega(v))$, где

$$\Psi(\delta, \Omega(v)) \equiv \psi_0(\delta, \|B\|\, g^{-1}(\Omega)) = \psi_0(\delta, \|B\| \cdot \|v\|).$$

Тогда вместо (2) можно использовать новое условие аппроксимации:

$$0 \leqslant [J_\delta(Bv) + \Psi(\delta, \Omega(v))] - J_0(Bv) \leqslant 2\Psi(\delta, \Omega(v)), \quad v \in \mathcal{V}. \tag{2'}$$

Из этого условия, в частности, получаются оценки

$$J_0^* = J_0(B\bar{v}) \leqslant J_0(Bv) \leqslant J_\delta(Bv) + \Psi(\delta, \Omega(v)), \tag{3}$$

$$|J_0^* - J_\delta(B\bar{v})| \leqslant \Psi(\delta, \bar{\Omega}). \tag{4}$$

Если ввести величину

$$\lambda_\delta \equiv \inf\{J_\delta(z) + \psi_0(\delta, \|z\|)\colon z \in D\},$$

то, основываясь на (2′), (3) и на оценке

$$\lambda_\delta \leqslant \inf\{J_\delta(Bv) + \Psi(\delta, \Omega(v))\colon v \in \mathcal{V}\},$$

можно установить неравенство

$$0 \leqslant \lambda_\delta - J_0^* \leqslant 2\Psi(\delta, \bar{\Omega}), \quad \bar{\Omega} \equiv \Omega(\bar{v}) > 0. \tag{5}$$

Фигурирующую в алгоритмах о. п. н., о. п. к., о. п. с. ф. величину I_δ введем на основании (3) следующим образом:

$$I_\delta(v)=f\left[J_\delta(Bv)+\Psi(\delta, \Omega(v))\right], \tag{6}$$

где вспомогательная функция $f(t)$ определена при $t\geqslant J_\delta^*$, неотрицательна, непрерывна, монотонно возрастает, подчиняясь условию $f(+\infty)=+\infty$, и, кроме того, является выпуклой.

Отметим, что условия применимости алгоритмов, сформулированные в общем виде в § 2, в данном частном случае оказываются выполненными.

Действительно, если в качестве топологии τ принять топологию слабой сходимости в V, то множество $\mathscr{V}$ рефлексивного банахова пространства V будет слабо замкнуто в силу условий на D и оператор B. На основании этого и с учетом слабой компактности шара в рефлексивном банаховом пространстве, а также принимая во внимание свойства функции $g(x)$, можно утверждать, что множества

$$\Omega_K=\{v\in\mathscr{V}\colon \Omega(v)\leqslant K\}=\{v\in\mathscr{V}\colon \|v\|\leqslant g^{-1}(K)\},\quad K>0$$

являются слабыми компактами в V (предположение 2 из § 2). Далее, исходя из слабой полунепрерывности снизу в V функционала $\|v\|$ и используя свойства непрерывности и монотонности функций ψ_0, g, можно утверждать, что функционалы $\Psi(\delta, \Omega(v))=\psi_0(\delta, \|B\|\cdot\|v\|)$, $\Omega(v)=g(\|v\|)$ слабо полунепрерывны снизу на D. Из данных о функционалах $J_0(z)$, $J_\delta(z)$ и из свойств оператора B легко сделать вывод о слабой полунепрерывности снизу на V функционалов $J_0(Bv)$, $J_\delta(Bv)$. Таким образом, функционал $J_\delta(Bv)+\Psi(\delta, \Omega(v))$ слабо полунепрерывен снизу на $\mathscr{V}$, и обосновано выполнение предположения 1 из § 2. Для того чтобы убедиться в справедливости предположения 3 из § 2, достаточно заметить, что функция Ψ, определяемая через функцию ψ_0, обладает всеми ее свойствами, в частности возрастает по второму аргументу.

Сформулируем теперь с учетом отмеченных особенностей постановки задачи алгоритмы о. п. н., о. п. к., о. п. с. ф. Для этого определим по данным $\{\Omega, I_\delta, \mathscr{V}\}$ сглаживающий функционал

$$M^\alpha[v]=\alpha\Omega(v)+I_\delta(v),\quad v\in\mathscr{V},\quad \alpha>0,$$

и рассмотрим экстремальную задачу: при заданном $\alpha>0$ и фиксированном δ найти элементы $v^\alpha\equiv v_\delta^\alpha\in\mathscr{V}$, для которых

$$M^\alpha[v^\alpha]=\inf\{M^\alpha[v]\colon v\in\mathscr{V}\}. \tag{7}$$

Как показано в общем случае в теореме 4.2, задача (7) разрешима при всяком $\alpha>0$. Обозначим символом $V^\alpha\equiv V_\delta^\alpha$ множество решений этой задачи при данном α и фиксированном δ. Тогда можно по аналогии с § 7—9 определить многозначные функции:

$$\begin{aligned}
&\beta(\alpha)\equiv\beta_\delta(\alpha)=I_\delta(v^\alpha),\\
&\gamma(\alpha)\equiv\gamma_\delta(\alpha)=\Omega(v^\alpha)\geqslant 0,\\
&\rho(\alpha)\equiv\rho_\delta(\alpha)=I_\delta(v^\alpha)-\Pi_\delta(v^\alpha),\\
&\varepsilon(\alpha)\equiv\varepsilon_\delta(\alpha)=M^\alpha[v^\alpha]-f\left[\lambda_\delta+2C\Psi(\delta, \Omega(v^\alpha))\right].
\end{aligned} \tag{8}$$

Здесь используется функция

$$\Pi_\delta(v^\alpha) \equiv f[\lambda_\delta + 2\Psi(\delta, \Omega(v^\alpha))],$$

а константа $C > 1$ фиксирована. Функции (8) обладают всеми свойствами, указанными в § 3, 7—9.

Алгоритм о. п. н. для решения экстремальной задачи (1), (1′), учитывающий априорную информацию о виде $\bar{z}_0$, заключается в нахождении приближенного решения задачи (1) в виде $z_\delta = Bv^{\alpha_\delta}$, где v^{α_δ} — специально выбранное решение задачи (7) при $\alpha = \alpha_\delta$. При этом параметр регуляризации $\alpha_\delta > 0$ выбирается как решение уравнения $\rho_\delta(\alpha) = 0$ с монотонной функцией $\rho_\delta(\alpha)$. Выбор экстремали v^{α_δ} и множества V^{α_δ} осуществляется по п р а в и л у о т б о р а 1. Пусть $C > 1$, $q > 1$ — заданные параметры, числа α_1, α_2 определены с помощью α_δ следующим образом: $\alpha_1 = \alpha_\delta / q$, $\alpha_2 = \alpha_\delta q$, а v^{α_1}, v^{α_2} — произвольные решения задачи (7) для $\alpha = \alpha_{1,2}$ соответственно. Если выполнено неравенство

$$I_\delta(v^{\alpha_2}) \geqslant C\Pi_\delta(v^{\alpha_1}) - f(\lambda_\delta)(C-1), \tag{9}$$

то в качестве v^{α_δ} выбирается такой элемент множества V^{α_δ}, для которого справедливо неравенство

$$I_\delta(v^{\alpha_\delta}) \leqslant \Pi_\delta(v^{\alpha_\delta}). \tag{10}$$

Если же

$$I_\delta(v^{\alpha_2}) \leqslant C\Pi_\delta(v^{\alpha_1}) - f(\lambda_\delta)(C-1), \tag{11}$$

то в качестве v^{α_δ} выбирается элемент множества V^{α_δ} такой, для которого справедливо неравенство

$$I_\delta(v^{\alpha_\delta}) \geqslant \Pi_\delta(v^{\alpha_\delta}). \tag{12}$$

Реализуемость такого алгоритма построения элемента v^{α_δ} установлена в § 7.

Перейдем к получению оценок для величин $\Omega(v^{\alpha_\delta})$, $J_\delta(Bv^{\alpha_\delta})$.

Т е о р е м а 1. *Для элемента v^{α_δ}, найденного по алгоритму о. п. н., справедлива оценка*

$$\Omega(v^{\alpha_\delta}) \leqslant \bar{\Omega}_1 \equiv \bar{\Omega}(C+q-1)/(C-1).$$

Д о к а з а т е л ь с т в о. Предположим, что элемент v^{α_δ} найден с помощью (9), (10) правила отбора 1. Если при этом выполнено неравенство

$$\alpha_2 \bar{\Omega} \geqslant (C-1)[f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)], \quad \bar{\Psi} \equiv \Psi(\delta, \bar{\Omega}),$$

то из него, а также из аналога

$$\alpha\Omega(v^\alpha) + I_\delta(v^\alpha) = M^\alpha[v^\alpha] \leqslant M^\alpha[\bar{v}] \leqslant \alpha\bar{\Omega} + f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})] \tag{13}$$

неравенства (5.2) получим

$$\Omega(v^{\alpha_\delta}) - \bar{\Omega} \leqslant [f(\lambda_\delta + 2\bar{\Psi}) - I_\delta(v^{\alpha_\delta})]/\alpha_\delta \leqslant$$
$$\leqslant q[f[\lambda_\delta + 2\bar{\Psi}] - f(\lambda_\delta)]/\alpha_2 \leqslant q\bar{\Omega}/(C-1). \tag{14}$$

Здесь использовано также вытекающее из определения числа λ_δ неравенство

$$f(\lambda_\delta) \leqslant f\left[J_\delta(Bv) + \Psi(\delta, \Omega(v))\right] \equiv I_\delta(v). \tag{15}$$

Из (14) следует доказываемое неравенство.

Если же оказывается, что

$$\alpha_2 \bar{\Omega} < (C-1)\left[f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)\right],$$

то из этого неравенства, из (9), (13) при $\alpha = \alpha_2$ с учетом возрастания функций f, Ψ, неотрицательности функции g, а также неубывания функции $\gamma(\alpha)$ получим

$$Cf\left[\lambda_\delta + 2\Psi(\delta, \Omega(v^{\alpha_\delta}))\right] \leqslant Cf\left[\lambda_\delta + 2\Psi(\delta, \Omega(v^{\alpha_1}))\right] = C\Pi_\delta(v^{\alpha_1}) \leqslant$$
$$\leqslant I_\delta(v^{\alpha_2}) + (C-1)f(\lambda_\delta) \leqslant M^{\alpha_2}\left[v^{\alpha_2}\right] + (C-1)f(\lambda_\delta) \leqslant$$
$$\leqslant \alpha_2\bar{\Omega} + f(\lambda_\delta + 2\bar{\Psi}) + (C-1)f(\lambda_\delta) \leqslant (C-1)\left[f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)\right] +$$
$$+ f(\lambda_\delta + 2\bar{\Psi}) + (C-1)f(\lambda_\delta) = Cf(\lambda_\delta + 2\bar{\Psi}).$$

Поскольку функция f является возрастающей, то из последней выкладки следует, что $\Psi(\delta, \Omega(v^{\alpha_\delta})) \leqslant \Psi(\delta, \bar{\Omega})$. Это неравенство в свою очередь с учетом возрастания Ψ по второму аргументу дает доказываемое неравенство $\Omega(v^{\alpha_\delta}) \leqslant \bar{\Omega} \leqslant \bar{\Omega}_1$.

Предположим теперь, что v^{α_δ} определено по второй части правила отбора 1. Тогда из (12) и оценки (13), взятой при $\alpha = \alpha_\delta$, получим

$$\alpha_\delta\Omega(v^{\alpha_\delta}) + f\left[\lambda_\delta + \Psi(\delta, \Omega(v^{\alpha_\delta}))\right] \leqslant \alpha_\delta\Omega(v^{\alpha_\delta}) + I_\delta(v^{\alpha_\delta}) =$$
$$= M^{\alpha_\delta}\left[v^{\alpha_\delta}\right] \leqslant \alpha_\delta\bar{\Omega} + f\left[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})\right],$$

т. е. $\Phi\left[\Omega(v^{\alpha_\delta})\right] \leqslant \Phi(\bar{\Omega})$. Здесь $\Phi(\Omega) \equiv f\left[\lambda_\delta + 2\Psi(\delta, \Omega)\right] + \alpha_\delta\Omega$. Из возрастания функций f, Ψ ясно, что функция $\Phi(\Omega)$ также возрастает при $\Omega \geqslant 0$. Поэтому $\Omega(v^{\alpha_\delta}) \leqslant \bar{\Omega} \leqslant \bar{\Omega}_1$. Теорема доказана.

Для получения оценки величины $J_\delta(Bv^{\alpha_\delta})$ потребуется

Лемма 1. *Если $f(t)$ — выпуклая при $t \geqslant a \geqslant 0$ функция, то для любого $x \geqslant 0$ справедливо неравенство*

$$Cf(a+x) - (C-1)f(a) \leqslant f(Cx+a), \quad C = \text{const} > 1.$$

Действительно, по определению выпуклой функции

$$f(a+x) = f\left[\frac{1}{C}(Cx+a) + \left(1 - \frac{1}{C}\right)a\right] \leqslant \frac{1}{C}f(Cx+a) + \left(1 - \frac{1}{C}\right)f(a),$$

откуда и следует утверждение леммы.

Теорема 2. *Для элемента $z_\delta = Bv^{\alpha_\delta}$, найденного по алгоритму о. п. н., выполняется оценка*

$$\lambda_\delta - \Psi(\delta, \bar{\Omega}_1) \leqslant J_\delta(z_\delta) \leqslant \lambda_\delta + 2C\Psi(\delta, \bar{\Omega}_1).$$

Доказательство. Оценка снизу вытекает из определения числа λ_δ и теоремы 1:

$$\lambda_\delta \leqslant J_\delta(z_\delta) + \Psi(\delta, \Omega(v^{\alpha_\delta})) \leqslant J_\delta(z_\delta) + \Psi(\delta, \bar{\Omega}_1).$$

Получим оценку сверху. В случае, если отбор элемента v^{α_δ} осуществляется по первой части правила отбора 1, то из (10) и по теореме 1 с учетом возрастания функций f, Ψ получаем

$$I_\delta(v^{\alpha_\delta}) \leqslant \Pi_\delta(v^{\alpha_\delta}) = f[\lambda_\delta + 2\Psi(\delta, \Omega(v^{\alpha_\delta}))] \leqslant f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega}_1)]. \quad (16)$$

В случае же, если отбор v^{α_δ} произведен по второй части правила отбора 1, рассмотрим два варианта. В первом выполнено неравенство

$$\Omega(v^{\alpha_1}) \leqslant \bar{\Omega}.$$

Тогда из (11), используя лемму 1 и неубывание функции $\beta(\alpha)$, получаем

$$I_\delta(v^{\alpha_\delta}) \leqslant I_\delta(v^{\alpha_2}) \leqslant Cf[\lambda_\delta + 2\Psi(\delta, \Omega(v^{\alpha_1}))] - (C-1) f(\lambda_\delta) \leqslant$$
$$\leqslant f[\lambda_\delta + 2C\Psi(\delta, \Omega(v^{\alpha_1}))] \leqslant f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega}_1)]. \quad (17)$$

Во втором варианте справедливо неравенство $\Omega(v^{\alpha_1}) > \bar{\Omega}$. Поэтому, если оказывается, что

$$\alpha_\delta \bar{\Omega} \geqslant (C-1)\{f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)\},$$

то, как в теореме 1, получим оценку $0 \leqslant \Omega(v^{\alpha_1}) - \bar{\Omega} \leqslant q\bar{\Omega}/(C-1)$. Таким образом, будет выполнено неравенство $\Omega(v^{\alpha_1}) \leqslant \bar{\Omega}_1$ и, применяя такие же рассуждения, как и в первом случае, можно снова получить (17). Если же верна оценка

$$\alpha_\delta \bar{\Omega} < (C-1)\{f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)\},$$

то из этого неравенства на основании (13) и леммы 1 следует

$$I_\delta(v^{\alpha_\delta}) \leqslant M^{\alpha_\delta}[v^{\alpha_\delta}] \leqslant \alpha_\delta \bar{\Omega} + f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})] \leqslant$$
$$\leqslant (C-1)\{f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})] - f(\lambda_\delta)\} + f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})] \leqslant$$
$$\leqslant Cf[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})] - (C-1) f(\lambda_\delta) \leqslant$$
$$\leqslant f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega})] \leqslant f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega}_1)]. \quad (18)$$

Итак, во всех возможных случаях из (16)—(18) получается, что

$$f[J_\delta(z_\delta)] \leqslant I_\delta(v^{\alpha_\delta}) \leqslant f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega}_1)]. \quad (19)$$

Отсюда, учитывая возрастание функции f, можно вывести искомую оценку.

Непосредственно из теоремы 2 и из (5) получается

Следствие 1. *Для приближенного решения $z_\delta = Bv^{\alpha_\delta}$ задачи* (1), *полученного по о. п. н., справедлива оценка скорости сходимости по функционалу*:

$$-\Psi(\delta, \bar{\Omega}_1) \leqslant J_\delta(z_\delta) - J_\delta^* \leqslant 2(C+1)\Psi(\delta, \bar{\Omega}_1).$$

Отметим также, что из результатов § 7 с учетом сказанного в § 11 о случае рефлексивного банахова пространства Z, наделенного топологией слабой сходимости, можно сделать вывод о слабой сходимости $v^{\alpha_\delta} \rightharpoonup \bar{v}$ и, значит, о сильной сходимости приближений $z_\delta \xrightarrow{z} \bar{z}_0$ при $\Delta \to 0$.

Рассмотрим теперь алгоритм о. п. к. для задачи (1), (1'). Он дает приближенное решение в виде $z_\delta = Bv^{\alpha_\delta}$, где $\alpha_\delta > 0$ и v^{α_δ} находятся следующим образом. Параметр регуляризации α_δ выбирается как решение уравнения с монотонной функцией $\gamma_\delta(\alpha) = \bar{\Omega}_\delta$, где $\bar{\Omega}_\delta$ — заданная оценка для $\bar{\Omega}$: $\bar{\Omega}_\delta \geqslant \bar{\Omega}$, $\bar{\Omega}_\delta \to \bar{\Omega}$ при $\Delta \to 0$ (см. § 8). Элемент v^{α_δ} находится по правилу отбора 2. Пусть C, q, α_1, α_2, v^{α_1}, v^{α_2} определяются по аналогии с величинами, используемыми в правиле отбора 1. Если выполнено неравенство (9), то в качестве v^{α_δ} выбирается такой элемент множества V^{α_δ}, для которого справедливо неравенство

$$\Omega\left(v^{\alpha_\delta}\right) \geqslant \bar{\Omega}_\delta; \tag{20}$$

если же выполнено (11), то v^{α_δ} выбирается из V^{α_δ} так, чтобы выполнялось неравенство

$$\Omega\left(v^{\alpha_\delta}\right) \leqslant \bar{\Omega}_\delta. \tag{21}$$

Реализуемость такого алгоритма построения приближенного решения задачи (1) обоснована в § 8.

Т е о р е м а 3. *Если число Δ столь мало, что выполнено неравенство $\bar{\Omega}_\delta \leqslant \bar{\Omega}_1$, то для элемента v^{α_δ}, найденного по алгоритму о. п. к., справедлива та же оценка, что и в теореме* 1.

Действительно, если v^{α_δ} определяется по первой части правила отбора 2, то искомая оценка получается, как для аналогичного случая в теореме 1 с использованием неравенств (9), (15), (13). Если же v^{α_δ} находится по второй части правила отбора 2, то при достаточно малых Δ выполнено неравенство $\bar{\Omega} \leqslant \bar{\Omega}_\delta \leqslant \bar{\Omega}_1$, и из (21) снова следует искомая оценка.

Т е о р е м а 4. *При выполнении условия теоремы* 3 *для приближенного решения $z_\delta = Bv^{\alpha_\delta}$ задачи* (1), *найденного по алгоритму о. п. к., справедлива та же оценка, что и в теореме* 2.

Д о к а з а т е л ь с т в о. Оценка снизу получается, как в теореме 2. При получении оценки сверху рассмотрим сначала случай отбора элемента v^{α_δ} по первой части правила отбора 2. Тогда из (13), (20) следует

$$\alpha_\delta \bar{\Omega} + I_\delta\left(v^{\alpha_\delta}\right) \leqslant \alpha_\delta \bar{\Omega}_\delta + I_\delta\left(v^{\alpha_\delta}\right) \leqslant \alpha_\delta \Omega\left(v^{\alpha_\delta}\right) + I_\delta\left(v^{\alpha_\delta}\right) \leqslant \alpha_\delta \bar{\Omega} + f\left[\lambda_\delta + 2\Psi\left(\delta, \bar{\Omega}\right)\right],$$

откуда получается (19). Если же отбор v^{α_δ} производится по второй части правила отбора 2, то можно применить рассуждения, аналогичные использованным в теореме 2 для второй части правила отбора 1. Эти рассуждения основываются на неравенствах (11), (13) и снова дают (19). Из соотношения (19) вытекает справедливость доказываемой оценки сверху.

С л е д с т в и е 2. *Если числа Δ таковы, что при $0 < \Delta \leqslant \Delta_0$ выполнено условие $\bar{\Omega}_\delta \leqslant \bar{\Omega}_1$, то для приближенного решения z_δ имеет место та же оценка, что и в следствии* 1.

Как и для алгоритма о. п. н., доказывается, что $z_\delta \xrightarrow{Z} \bar{z}_0$ при $\Delta \to 0$.

Обратимся, наконец, к алгоритму о. п. с. ф. Для задачи (1), (1′) он определяет приближенное решение в виде $z_\delta = Bv^{\alpha_\delta}$, где параметр $\alpha_\delta > 0$ выбирается как решение уравнения $\varepsilon(\alpha) = 0$ с монотонной функцией $\varepsilon(\alpha)$ (см. § 9), а элемент v^{α_δ} находится по правилу отбора 3. В качестве v^{α_δ} берется такой элемент из множества V^{α_δ}, для которого выполнено неравенство

$$M^{\alpha_\delta}[v^{\alpha_\delta}] \geqslant f[\lambda_\delta + 2C\Psi(\delta, \Omega(v^{\alpha_\delta}))]. \tag{22}$$

Возможность реализации такого алгоритма устанавливается, как и в § 9. Данный вариант алгоритма о.п.с.ф. отличается от изложенного в § 9 заменой членов Ψ^p на члены $2C\Psi$. Как будет показано ниже, такой вариант алгоритма обеспечивает сходимость приближенных решений z_δ к единственному решению $\bar{z}$ задачи (1).

Теорема 5. *Для элементов v^{α_δ}, найденных по алгоритму о.п.с.ф., выполнена оценка*

$$\Omega(v^{\alpha_\delta}) \leqslant \bar{\Omega}_2 \equiv C\bar{\Omega}/(C-1).$$

Доказательство. Если v^{α_δ} удовлетворяет неравенству $\Omega(v^{\alpha_\delta}) \leqslant \bar{\Omega}$, то искомая оценка очевидным образом выполняется. Пусть теперь $\Omega(v^{\alpha_\delta}) > \bar{\Omega}$. Тогда из этого неравенства, из условия (22) и из (13) с учетом возрастания функций f, Ψ можно получить

$$f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega})] \leqslant Cf[\lambda_\delta + 2C\Psi(\delta, \Omega(v^{\alpha_\delta}))] \leqslant \\ \leqslant M^{\alpha_\delta}[v^{\alpha_\delta}] \leqslant \alpha_\delta \bar{\Omega} + f[\lambda_\delta + 2\Psi(\delta, \bar{\Omega})], \tag{23}$$

$$0 < \Omega(v^{\alpha_\delta}) - \bar{\Omega} \leqslant [f[\lambda_\delta + 2\bar{\Psi}] - f(\lambda_\delta)]/\alpha_\delta. \tag{24}$$

Из (23) ясно, что

$$\alpha_\delta \geqslant \bar{\Omega}^{-1}\{f(\lambda_\delta + 2C\bar{\Psi}) - f(\lambda_\delta + 2\bar{\Psi})\},$$

и поэтому из (24) следует соотношение

$$0 < \Omega(v^{\alpha_\delta}) - \bar{\Omega} \leqslant \bar{\Omega}[f(\lambda_\delta + 2\bar{\Psi}) - f(\lambda_\delta)]/[f(\lambda_\delta + 2C\bar{\Psi}) - \\ - f(\lambda_\delta + 2\bar{\Psi})] \leqslant \bar{\Omega}/(C-1).$$

Последнее неравенство в этой оценке получено на основании леммы 1.

Теорема 6. *Для приближенного решения $z_\delta = Bv^{\alpha_\delta}$, полученного по алгоритму о.п.с.ф., имеет место оценка*

$$\lambda_\delta - \Psi(\delta, \bar{\Omega}_2) \leqslant J_\delta(z_\delta) \leqslant \lambda_\delta + 2C\Psi(\delta, \bar{\Omega}_2).$$

Доказательство. Рассмотрим два случая. В первом считается, что выполнено неравенство $\Omega(v^{\alpha_1}) \leqslant \bar{\Omega}$, где $\alpha_1 \equiv \alpha_\delta/q$ $(q > 1)$, v^{α_1} — некоторое решение задачи (7) при $\alpha = \alpha_1$. Тогда вследствие неубывания функции $\varepsilon(\alpha)$ и по условию выбора параметра α_δ оказывается, что

$$\varepsilon(\alpha_1) = M^{\alpha_1}[v^{\alpha_1}] - f[\lambda_\delta + 2C\Psi(\delta, \Omega(v^{\alpha_1}))] \leqslant \varepsilon(\alpha_\delta - 0) \leqslant 0$$

(см. следствие 6.5). Отсюда и из неравенства для $\Omega(v^{\alpha_1})$ получим

$$\alpha_1 \Omega(v^{\alpha_1}) + I_\delta(v^{\alpha_1}) \leqslant f[\lambda_\delta + 2C\Psi(\delta, \bar{\Omega})].$$

В свою очередь из этого неравенства вследствие экстремальности элемента v^{α_δ} как решения задачи (7) при $\alpha=\alpha_\delta$ получается соотношение

$$I_\delta(v^{\alpha_\delta})\leqslant M^{\alpha_\delta}[v^{\alpha_\delta}]\leqslant M^{\alpha_\delta}[v^{\alpha_1}]=\alpha_\delta\Omega(v^{\alpha_1})+I_\delta(v^{\alpha_1})\leqslant$$
$$\leqslant q[\alpha_1\Omega(v^{\alpha_1})+I_\delta(v^{\alpha_1})]\leqslant qf[\lambda_\delta+2C\bar\Psi(\delta,\bar\Omega)].$$

Таким образом,

$$I_\delta(v^{\alpha_\delta})\leqslant qf[\lambda_\delta+2C\Psi(\delta,\bar\Omega)]\leqslant qf[\lambda_\delta+2C\Psi(\delta,\bar\Omega_2)]. \tag{25}$$

Во втором случае считается выполненным неравенство $\Omega(v^{\alpha_1})>\bar\Omega$. Если при этом оказывается, что $\alpha_\delta\bar\Omega\geqslant(C-1)[f(\lambda_\delta+2\bar\Psi)-f(\lambda_\delta)]$, то, как при доказательстве аналогичной части теоремы 2, можно найти оценку $\Omega(v^{\alpha_1})\leqslant\bar\Omega(C+q-1)/(C-1)$. Из нее, как и в первом случае, выводим аналог соотношения (25):

$$I_\delta(v^{\alpha_\delta})\leqslant qf[\lambda_\delta+2C\Psi(\delta,\bar\Omega(C+q-1)/(C-1))]. \tag{25'}$$

Если же выполнено неравенство $\alpha_\delta\bar\Omega<(C-1)[f(\lambda_\delta+2\bar\Psi)-f(\lambda_\delta)]$, то, применяя методику аналогичного случая теоремы 2 (см. (18)), снова получим оценку (25).

Итак, во всех рассмотренных случаях справедливы неравенства (25) или (25′). Из них с учетом произвольности числа $q>1$ и вследствие непрерывности функций f, Ψ легко выводится оценка

$$I_\delta(v^{\alpha_\delta})\leqslant f[\lambda_\delta+2C\Psi(\delta,\bar\Omega_2)].$$

Это неравенство, являясь аналогом соотношения (19), как и в теореме 2, дает искомую верхнюю оценку. Нижняя оценка для $J_\delta(z_\delta)$ устанавливается, как в теоремах 2, 4.

Следствие 3. *Для приближенного решения z_δ, получаемого по алгоритму о.п.с.ф., справедлива оценка скорости сходимости по функционалу*

$$-\Psi(\delta,\bar\Omega_2)\leqslant J_\delta(z_\delta)-J_0^*\leqslant 2(C+1)\Psi(\delta,\bar\Omega_2).$$

Изучим вопрос о сходимости алгоритма о.п.с.ф. Для этого понадобится следующая вспомогательная

Лемма 2. *Предположим, что в рамках данной постановки задачи выполнены условия*:

а) *функционал $J_0(z)$ полунепрерывен снизу на замкнутом в Z множестве D*;

б) *задача* (1′) *имеет единственное решение $\bar v\in\mathscr{V}$*;

в) *семейство элементов $\{v_\delta\}\subset\mathscr{V}$ ($\|\delta\|\leqslant\Delta_0$) равномерно ограничено по норме константой C_0: $\|v_\delta\|\leqslant C_0$*;

г) $J_\delta(Bv_\delta)\to J_0^*$ *при* $\Delta\to0$.

Тогда v_δ слабо сходится к $\bar v$ при $\Delta\to0$, а $z_\delta\equiv Bv_\delta$ сильно сходится к $\bar z=B\bar v$ при $\Delta\to0$.

Доказательство. Рассмотрим произвольную последовательность $\{v_n\}$, где $v_n\equiv v_{\delta_n}$, соответствующую последовательности $\{\delta_n\}$, сходящейся при $n\to\infty$ к нулю. По условию в) эта последовательность будет слабо компактной на слабо замкнутом множестве $\mathscr{V}$ рефлексивного банахова пространства V (см. теорему 0.3.6). Поэтому из $\{v_n\}$

можно выбрать слабо сходящуюся последовательность $\{v_{n_k}\}: v_{n_k} \rightharpoonup v_0$ при $k\to\infty$. Предел v_0 принадлежит множеству $\mathscr{V}$ вследствие слабой замкнутости последнего. Но тогда из полной непрерывности оператора B вытекает сильная сходимость: $z_{n_k}\equiv Bv_{n_k}\xrightarrow{Z} z_0\equiv Bv_0\in D$.

Далее из условия аппроксимации (2) и условия в) леммы следует, что

$$J_0(z_{n_k})\leqslant J_{\delta_{n_k}}(z_{n_k})+\psi_0(\delta_{n_k}, \|z_{n_k}\|)=J_{\delta_{n_k}}(Bv_{n_k})+\psi_0(\delta_{n_k}, \|Bv_{n_k}\|)\leqslant$$
$$\leqslant J_{\delta_{n_k}}(Bv_{n_k})+\psi_0(\delta_{n_k}, \|B\|\cdot C_0).$$

Отсюда, переходя к пределу при $k\to\infty$ и учитывая условия а), г) леммы, а также свойства функции ψ_0, получаем

$$J_0^*\leqslant J_0(z_0)\leqslant \varliminf_{k\to\infty} J_0(z_{n_k})\leqslant \varliminf_{k\to\infty}\left[J_{\delta_{n_k}}(Bv_{n_k})+\psi_0(\delta_{n_k}, \|B\|\cdot C_0)\right]=J_0^*.$$

Таким образом, $J_0(z_0)=J_0(Bv_0)=J_0^*$. Вследствие единственности решения задачи (1′) $v_0=\bar v$, так что $v_{n_k}\rightharpoonup\bar v$ при $k\to\infty$.

Покажем теперь, что вся последовательность $\{v_n\}$ слабо сходится к $\bar v$ при $n\to\infty$. Пусть это не так. Тогда в силу относительно слабой компактности множества $\{v_n\}$ найдется подпоследовательность $\{v_{n_p}\}\subset\{v_n\}$, которая слабо сходится при $p\to\infty$ к $v^*\in\mathscr{V}$, причем $v^*\neq\bar v$. Повторяя для подпоследовательности $\{v_{n_p}\}$ все рассуждения, которые были применены для $\{v_{n_k}\}$, можно прийти к выводу, что $v^*=\bar v$. Это противоречие доказывает сходимость $v_n\rightharpoonup\bar v$ при $n\to\infty$. Сильная сходимость $z_n=Bv_n\to B\bar v=\bar z$ вытекает из полной непрерывности оператора B.

Теорема 7. *При $\Delta\to 0$ для элементов $v^{\alpha\delta}$, $z_\delta=Bv^{\alpha\delta}$, полученных по о.п.с.ф., выполнены предельные соотношения $v^{\alpha\delta}\rightharpoonup\bar v$, $z_\delta\xrightarrow{Z}\bar z_0$.*

Доказательство. Заметим, что, используя свойства функции g, можно получить из теоремы 5 оценку

$$\|v^{\alpha\delta}\|\leqslant g^{-1}(\bar\Omega_2)\equiv C_0=\text{const}\quad \forall\delta,\ \|\delta\|\leqslant\Delta_0,$$

а из следствия 3 сходимости $J_\delta(z_\delta)=J_\delta(Bv^{\alpha\delta})\to J_0^*$ при $\Delta\to 0$. Тогда для $\{v^{\alpha\delta}\}$ выполнены условия леммы 2, из которой и получаются доказываемые утверждения.

Необходимо подчеркнуть, что оценки из теорем 1—6 получены без использования свойства полной непрерывности оператора B и поэтому, в частности, справедливы и для случая, когда $V=Z$, а B — единичный оператор из Z в Z. Полная непрерывность оператора B использовалась до сих пор лишь для обоснования сильной сходимости приближений.

Перейдем теперь к исследованию точности алгоритмов о.п.н., о.п.к., о.п.с.ф. в сравнении с другими методами решения задачи (1), (1′).

Определение 1. Назовем *методом решения задачи* (1) оператор P, сопоставляющий каждому набору данных $(J_\delta, \delta, \psi_0)$ элемент $z^\delta=P(J_\delta, \delta, \psi_0)\in D$.

Как уже отмечалось в начале параграфа, спецификой некорректно поставленных задач является невозможность получения оценок точности методов их решения без привлечения дополнительной информации о свойствах точных решений задачи. Наиболее употребительной информацией такого рода является принадлежность некоторому заданному компакту. В связи с этим для получения оценок точности методов P решения задачи (1), (1′) будем считать, что справедливо неравенство $\|\bar{v}\| \leqslant r$, где r — фиксированное число. Тогда решение $\bar{z}_0$ задачи (1) принадлежит множеству корректности [88, 110, 189] этой задачи $M_r = \{z \in D: z = Bv, \ \|v\| \leqslant r\}$.

Отметим, что из условия аппроксимации (2) на множестве следует условие равномерной аппроксимации на множестве M_r:

$$|J_0(z) - J_\delta(z)| \leqslant \psi_0(\delta, \|B\| r) \quad \forall z \in M_r.$$

Рассмотрим множество всех задач типа (1): найти элемент $\bar{z} \in D$ такой, что

$$J(\bar{z}) = \inf\{J(z): z \in D\}, \tag{26}$$

в которых фигурирует полунепрерывный снизу на D точный функционал J и которые имеют единственное решение $\bar{z}$, причем $\bar{z} \equiv \bar{z}(J) = B\bar{w}$ $(\bar{w} \in V)$.

Пусть $R \equiv R(r) \geqslant r$ — некоторое число, вид которого будет указан ниже. Обозначим через $\{J\}_R$ множество функционалов из задач (26), для которых $\bar{z}(J) \in M_R$. Такие задачи вида (26) эквивалентны следующим задачам:

$$J(\bar{z}) = \inf\{J(z): z \in M_R\}. \tag{27}$$

Выделим теперь из $\{J\}_R$ при фиксированных приближенных данных $\{J_\delta, \delta, \psi_0\}$ подмножество A_R точных функционалов J:

$$A_R \equiv A_R(J_\delta, \delta, \psi_0) \equiv A_R(J_\delta, \delta) =$$
$$= \{J: J \in \{J\}_R, \ |J(\bar{z}) - \lambda_\delta| \leqslant 2H, \ |J(z) - J_\delta(z)| \leqslant 2CH \ \forall z \in M_R\}.$$

Здесь $\lambda_\delta = \lambda_\delta(J_\delta, \delta)$ — число, определенное в начале параграфа, $H \equiv \psi_0(\delta, \|B\| R)$, $C > 1$ — параметр из алгоритмов о.п.н., о.п.к., о.п.с.ф. Согласно неравенствам (2), (5), выполнено включение $J_0 \in A_R(J_\delta, \delta)$. Множество A_R задает определенный класс задач (27), для которых функционал J_δ может служить приближенным функционалом в смысле усиленного условия равномерной аппроксимации:

$$|J(z) - J_\delta(z)| \leqslant 2CH \quad \forall z \in M_R.$$

Будем характеризовать точность метода P решения задач (27) из класса $A_R(J_\delta, \delta)$ с приближенными данными (J_δ, δ) величиной

$$\Delta_P(J_\delta, \delta) = \sup_J \{\|P(J_\delta, \delta) - \bar{z}(J)\|: J \in A_R(J_\delta, \delta)\}.$$

Определение 2. *Оптимальной точностью приближенного решения задач* (27) *на данных* (J_δ, δ) назовем число

$$\Delta_{\text{опт}} = \Delta_{\text{опт}}(J_\delta, \delta) = \inf\{\Delta_P(J_\delta, \delta): P \in \mathscr{P}\},$$

где $\mathscr{P}$ — класс всех возможных методов решения задач вида (27).

Определение 3. Будем говорить, что метод P_0 решения задачи (1) имеет *оптимальный порядок точности на данных* (J_δ, δ) (*является оптимальным по порядку*), если

$$\|P_0(J_\delta, \delta)-\bar{z}_0\| \leqslant k\Delta_{\text{опт}}$$

при $\|\delta\| \leqslant \Delta_0=\text{const}$, причем число $k>0$ не зависит от J_δ, δ, R.

В дальнейшем будет также использоваться оценочная функция

$$\Theta(H, R)\equiv\Theta(H, R, J_\delta, \delta)=\sup_{J_1, J_2}\{\|z(J_1)-z(J_2)\|: J_{1,2}\in A_R(J_\delta, \delta)\},$$

определяемая при фиксированных R, J_δ, δ. Функция Θ обладает важным свойством.

Лемма 3. *Справедлива оценка* $\Delta_{\text{опт}} \geqslant \Theta(H, R)/2$.

Доказательство. По определению точной верхней грани для каждого $\varepsilon>0$ найдутся такие элементы J_1, $J_2\in A_R(J_\delta, \delta)$, для которых выполнено неравенство $\|z(J_1)-z(J_2)\| \geqslant \Theta(H, R)-\varepsilon$. Тогда

$$\Delta_P(J_\delta, \delta) \geqslant \max\{\|P(J_\delta, \delta)-\bar{z}(J_1)\|,\ \|P(J_\delta, \delta)-\bar{z}(J_2)\|\} \geqslant$$
$$\geqslant \|\bar{z}(J_1)-\bar{z}(J_2)\|/2 \geqslant \Theta(H, R)/2-\varepsilon \quad \forall P\in\mathscr{P}.$$

Из этого соотношения в силу произвольности $\varepsilon>0$ и $P\in\mathscr{P}$ получается искомая оценка.

Введем теперь множество $Z^\delta(\tau, R)=\{z\in M_R: J_\delta(z)-\lambda_\delta \leqslant \tau\}$. Это множество непусто при $\tau \geqslant H$, так как из неравенств (4), (5) следует $J_\delta(\bar{z}_0)-\lambda_\delta \leqslant \Psi(\delta, \bar{\Omega}) \leqslant H$ и, значит, $\bar{z}_0\in Z^\delta(\tau, R)$. Определим далее оценочную функцию, непосредственно связанную с (J_δ, δ):

$$\omega_\delta(\tau, R)=\sup\{\|z_1-z_2\|: z_{1,2}\in Z^\delta(\tau, R)\}, \quad \tau \geqslant H. \tag{28}$$

Связь Θ и ω_δ устанавливает

Лемма 4. $\Theta(H, R) \geqslant \omega_\delta(2CH, R)$.

Доказательство. Вследствие компактности множества M_R в Z и по определению величины $\omega_\delta(2CH, R)$ найдутся такие элементы z_1, $z_2\in M_R$, для которых выполнены соотношения

$$\|z_1-z_2\|=\omega_\delta(2CH, R), \quad J_\delta(z_{1,2})-\lambda_\delta \leqslant 2CH. \tag{29}$$

Введем вспомогательные функционалы:

$$J_1(z)=\{\lambda_\delta: z=z_1;\ J_\delta(z)+2CH: z\neq z_1\},$$
$$J_2(z)=\{\lambda_\delta: z=z_2;\ J_\delta(z)+2CH: z\neq z_2\}.$$

В силу неравенства

$$\lambda_\delta=J_{1,2}(z_{1,2}) \leqslant J_\delta(z)+\psi_0(\delta, \|z\|)<J_\delta(z)+2C\psi_0(\delta, \|B\|\, R)=$$
$$=J_{1,2}(z) \quad \forall z\in M_R, \quad z\neq z_{1,2},$$

вытекающего из определения числа λ_δ и условия $C>1$, можно сделать вывод, что $z_{1,2}=\bar{z}(J_{1,2})$ и что функционалы $J_{1,2}$ полунепрерывны снизу на D. Очевидно, что $|J_{1,2}(z_{1,2})-\lambda_\delta|=0<2H$. Кроме того,

$$|J_1(z)-J_\delta(z)|=\{|J_\delta(z)+2CH-J_\delta(z)|: z\neq z_1;\ |\lambda_\delta-J_\delta(z_1)|: z=z_1\} \leqslant 2CH,$$

где использовано неравенство (29). Аналогичным образом

$$|J_2(z)-J_\delta(z)|\leqslant 2CH.$$

Поэтому $J_{1,2}(z)\in A_R(J_\delta,\ \delta)$; следовательно, по определению функции $\Theta(H,\ R)$, согласно равенству (29), оказывается, что

$$\Theta(H,\ R)\geqslant\|\bar{z}(J_1)-\bar{z}(J_2)\|=\|z_1-z_2\|=\omega_\delta(2CH,\ R).$$

Следствие 4. $\Delta_{\text{опт}}\geqslant\omega_\delta(2CH,\ R)/2$.

Установим некоторые дополнительные свойства оценочной функции (28). Для этого введем

$$\omega_0(\tau,\ R)=\sup\{\|z_1-z_2\|:\ z_{1,2}\in M_R,\ J_0(z_{1,2})-J_0^*\leqslant\tau\}$$

и исследуем связь $\omega_0(\tau,\ R)$ и $\omega_\delta(\tau,\ R)$.

Лемма 5. *Для каждой константы* $p\geqslant 1$ *выполнено неравенство* $\omega_\delta(pH,\ R)\leqslant\omega_0((p+3)H,\ R)$.

Доказательство. Рассмотрим произвольные элементы $z_{1,2}\in M_R$, для которых $J_\delta(z_{1,2})-\lambda_\delta\leqslant pH$. Поскольку $\|z_{1,2}\|\leqslant\|B\|\ R$, то вследствие неравенств (2), (5)

$$0\leqslant J_0(z_{1,2})-J_0^*\leqslant|J_0(z_{1,2})-J_\delta(z_{1,2})|+[J_\delta(z_{1,2})-\lambda_\delta]+|\lambda_\delta-J_0^*|\leqslant$$
$$\leqslant\psi_0(\delta,\ \|z_{1,2}\|)+pH+2\psi_0(\delta,\ \|B\|\ R)\leqslant 3\psi_0(\delta,\ \|B\|\ R)+pH\equiv(p+3)H.$$

Тем самым по определению функции $\omega_0(\tau,\ R)$ оказывается, что $\|z_1-z_2\|\leqslant\omega_0\ [(p+3)H,\ R]$, откуда по определению функции $\omega_\delta(\tau, R)$ и вследствие отмеченной произвольности элементов $z_1,\ z_2$ следует искомая оценка.

Большой интерес представляет изучение поведения функций $\omega_0(pH,\ R)$, $\omega_\delta(pH,\ R)$ при $\Delta\to 0$ $(H\to 0)$.

Лемма 6. *Из условия единственности решения задачи* (1) *следует, что для всякой константы* $p\geqslant 1$ *имеет место сходимость* $\omega_0(pH,\ R)\to 0$, $\omega_\delta(pH,\ R)\to 0$ *при* $\Delta\to 0$.

Доказательство. Возьмем произвольную последовательность $\{\delta_n\}$ такую, что $\Delta_n=\|\delta_n\|\to 0$ при $n\to\infty$. Тогда $H_n=\psi_0(\delta_n,\ \|B\|\ R)\to 0$. Предположим, что лемма неверна и $\omega_0(pH_n,\ R)\geqslant\varepsilon_0=\text{const}>0$ для любого номера n. По определению функции ω_0 и в силу компактности множества M_R для каждого n найдутся такие элементы $z_{1,2}^{(n)}\in M_R$, для которых

$$\omega_0(pH_n,\ R)=\|z_1^{(n)}-z_2^{(n)}\|\geqslant\varepsilon_0>0,\quad 0\leqslant J_0(z_{1,2}^{(n)})-J_0^*\leqslant pH_n. \tag{30}$$

Выберем из $\{z_{1,2}^{(n)}\}$ сходящиеся подпоследовательности $z_1^{(n_l)}\to z_1$, $z_2^{(n_l)}\to z_2$ при $l\to\infty$, $z_{1,2}\in M_R\subset D$. Тогда из (30), учитывая полунепрерывность снизу на D функционала J_0, получаем

$$\begin{gathered}\lim_{l\to\infty}\omega_0(pH_{n_l},\ R)=\lim_{l\to\infty}\|z_1^{(n_l)}-z_2^{(n_l)}\|=\|z_1-z_2\|\geqslant\varepsilon_0>0,\\ 0\leqslant J_0(z_{1,2})-J_0^*\leqslant\varliminf_{l\to\infty}J_0(z_{1,2}^{(n_l)})-J_0^*\leqslant\lim_{l\to\infty}pH_{n_l}=0.\end{gathered} \tag{31}$$

Поэтому $J_0(z_{1,2})=J_0^*$, и вследствие единственности решения задачи (1) оказывается, что $z_1=z_2$. Это, однако, противоречит неравенству

(31). Поэтому $\omega_0(pH_n, R)\to 0$ при $\Delta_n\to 0$. Из леммы 5 тогда ясно, что и $\omega_{\delta_n}(pH_n, R)\to 0$ при $\Delta_n\to 0$.

Лемма 7. *Из условия единственности решения задачи* (1) *следует, что* $\Delta_{\text{опт}}(J_\delta, \delta)\to 0$ *при* $\Delta\to 0$.

Доказательство. Рассмотрим специальный метод P^0 решения задач (27) с функционалами $J\in A_R(J_\delta, \delta)$: $P^0(J_\delta, \delta)=z_\delta^0$. Здесь z_δ^0 — некоторый элемент множества $Z^\delta(4CH, R)\equiv\{z\in M_R: J_\delta(z)-\lambda_\delta\leqslant 4CH\}$. Поскольку $J\in A_R(J_\delta, \delta)$, то $\bar z=\bar z(J)\in M_R$ и

$$J_\delta[\bar z(J)]-\lambda_\delta\leqslant|J_\delta(\bar z)-J(\bar z)|+|J(\bar z)-\lambda_\delta|\leqslant 2CH+2H\leqslant 4CH.$$

Тогда по определению величины $\omega_\delta(4CH, R)$ (см. (28)) выполнено неравенство

$$\|P^0(J_\delta, \delta)-\bar z(J)\|\equiv\|z_\delta^0-\bar z\|\leqslant\omega_\delta(4CH, R).$$

Правая часть этой оценки не зависит от $J\in A_R(J_\delta, \delta)$, так что

$$\Delta_{\text{опт}}\leqslant\Delta_{P^0}(J_\delta, \delta)=\sup\{\|P^0(J_\delta, \delta)-\bar z(J)\|: J\in A_R(J_\delta, \delta)\}\leqslant\omega_\delta(4CH, R),$$

и по лемме 6 справедливо доказываемое предельное соотношение.

Из леммы 7 можно сделать вывод: величина $\|P(J_\delta, \delta)-\bar z_0\|$, которая характеризует точность приближенного решения задачи (1), полученного с помощью метода P, стремится к нулю при $\Delta\to 0$ для оптимальных по порядку точности методов.

Теорема 8. *Алгоритмы о.п.н., о.п.к. и о.п.с.ф. имеют оптимальный порядок точности на данных* (J_δ, δ).

Доказательство. Определим число $R(r)$, положив $R(r)=(C+q-1)r/(C-1)$ для о.п.н., о.п.к. и $R(r)=Cr/(C-1)$ для о.п.с.ф. Тогда, как следует из теорем 1—6 для приближенных решений $z_\delta(J_\delta, \delta)=Bv^{\alpha_\delta}$, полученных по этим алгоритмам, выполнены оценки

$$\begin{gathered} g(\|v^{\alpha_\delta}\|)=\Omega(v^{\alpha_\delta})\leqslant g(R), \\ |J_\delta(z_\delta)-\lambda_\delta|\leqslant 2C\Psi(\delta, g(R))=2C\psi_0(\delta, \|B\|\,R)\equiv 2CH. \end{gathered} \tag{32}$$

Отсюда, в частности, ясно, что $z_\delta\in Z^\delta(2CH, R)$. Кроме того, по предположению $\|\bar v\|\leqslant r<R$, а вследствие (4), (5)

$$-H\leqslant-\psi_0(\delta, \|\bar z_0\|)\leqslant J_\delta(\bar z_0)-\lambda_\delta\leqslant J_\delta(\bar z_0)-\lambda_0\leqslant\psi_0(\delta, \|\bar z_0\|)\leqslant H.$$

Поэтому $\bar z_0\in Z^\delta(H, R)\subset Z^\delta(2CH, R)$. Тогда из определения (28) и следствия 4 легко получается неравенство

$$\|z_\delta(J_\delta, \delta)-\bar z_0\|\leqslant\omega_\delta(2CH, R)\leqslant 2\Delta_{\text{опт}},$$

которое и доказывает теорему.

Рассмотрим теперь варианты алгоритмов о.м.н. и м.к. для задачи (1). В них в качестве приближения к $\bar z_0$ берется элемент $z_\delta=Bv_\delta$. При этом $v_\delta\in\mathscr{V}$ находится следующим образом: для о.м.н. v_δ есть любое решение экстремальной задачи

$$\Omega(v_\delta)=\inf\{\Omega(v): v\in\mathscr{V},\ J_\delta(Bv)-\lambda_\delta\leqslant\Psi(\delta, \Omega(v))\}, \tag{33}$$

для м.к. v_δ есть любое решение задачи

$$J_\delta(Bv)=\inf\{J_\delta(Bv): v\in\mathscr{V},\ \Omega(v)\leqslant\bar\Omega_\delta\}. \tag{34}$$

Существование таких элементов v_δ следует из результатов § 12. Дадим оценки функционалов от v_δ, аналогичные приведенным в теоремах 1—6.

Теорема 9. *Пусть v_δ, $\tilde{v}_\delta$ получены по алгоритмам о.м.н., м.к. соответственно. Тогда при* $0<\|\delta\|\leqslant\Delta_0=\text{const}$

$$\Omega(v_\delta)\leqslant\bar{\Omega},\quad \Omega(\tilde{v}_\delta)\leqslant\bar{\Omega}_0\equiv C\bar{\Omega},\qquad C>1, \tag{35}$$

$$|J_\delta(Bv_\delta)-\lambda_\delta|\leqslant\Psi(\delta,\bar{\Omega}),\qquad |J_\delta(B\tilde{v}_\delta)-\lambda_\delta|\leqslant\Psi(\delta,\bar{\Omega}_0). \tag{36}$$

Доказательство. Возьмем такое Δ_0, что при $\|\delta\|\leqslant\Delta_0$ выполнено неравенство $\bar{\Omega}_\delta\leqslant\bar{\Omega}_0$. Тогда неравенства (35) получаются из формулировок задач (33), (34). Оценки (35) и определение числа λ_δ дают неравенства

$$\begin{aligned}\lambda_\delta&\leqslant J_\delta(Bv_\delta)+\Psi(\delta,\Omega(v_\delta))\leqslant J_\delta(Bv_\delta)+\Psi(\delta,\bar{\Omega}),\\ \lambda_\delta&\leqslant J_\delta(B\tilde{v}_\delta)+\Psi(\delta,\Omega(\tilde{v}_\delta))\leqslant J_\delta(B\tilde{v}_\delta)+\Psi(\delta,\bar{\Omega}_0).\end{aligned} \tag{37}$$

Из определения элемента $v_\delta\in\mathscr{V}$ ясно, что

$$J_\delta(Bv_\delta)\leqslant\lambda_\delta+\Psi(\delta,\Omega(v_\delta))\leqslant\lambda_\delta+\Psi(\delta,\bar{\Omega}), \tag{38}$$

а из задачи (34) и неравенства $\lambda_\delta\geqslant J_0^*$ получим

$$J_\delta(B\tilde{v}_\delta)-\lambda_\delta\leqslant J_\delta(B\bar{v})-J_0^*=J_\delta(B\bar{v})-J_0(B\bar{v})\leqslant\Psi(\delta,\bar{\Omega})\leqslant\Psi(\delta,\bar{\Omega}_0). \tag{39}$$

Здесь использовано также соотношение (4). Найденные неравенства (37)—(39) доказывают справедливость оценок (36).

Следствие 5. *Алгоритмы о.м.н., м.к. имеют оптимальный порядок точности на данных* (J_δ,δ).

Доказательство проводится по той же схеме, что и в теореме 7. При этом принимается $R(r)=r$ для о.м.н. и $R(r)=Cr$ для м.к., так что соотношения (35), (36) становятся аналогами оценок (32) из теоремы 7:

$$\Omega(v_\delta)\leqslant g(R),\qquad \Omega(\tilde{v}_\delta)\leqslant g(R),$$
$$|J_\delta(Bv_\delta)-\lambda_\delta|\leqslant\Psi(\delta,g(R))\leqslant 2CH,\qquad |J_\delta(B\tilde{v}_\delta)-\lambda_\delta|\leqslant 2CH.$$

Замечание. Вопросы оптимальности по порядку точности алгоритмов решения экстремальных задач рассматривались также в [34, 142]. В частности, в [142] рассмотрено свойство оптимальности по порядку точности методов невязки и квазирешений при специальной (равномерной) аппроксимации точного функционала приближенным, т. е. в случае $\psi_0(\delta,\|z\|)=\psi_0(\delta)$.

§ 14. Алгоритмические особенности методов

Реализация алгоритмов из § 7—9 связана с необходимостью эффективно решать следующие вспомогательные задачи.

1. Нахождение величины $\lambda_\delta=\inf\{J_\delta(z)+\Psi(\delta,\Omega(z)): z\in D\}$.

2. Решение экстремальной задачи (4.2) минимизации сглаживающего функционала при фиксированных $\alpha>0$ и вычисление на этой основе функций (7.2), (8.1), (9.1), которые используются для выбора параметра регуляризации.

3. Нахождение какого-либо решения α_δ одного из уравнений (7.3), (8.1), (9.2).

4. Отбор экстремали z^{α_δ} из их множества, отвечающего найденному параметру регуляризации α_δ по соответствующим правилам (см. 7.1, 8.1, 9.1).

Этот набор задач 1—4 представляет собой схему реализации алгоритмов о.п.н., о.п.к., о.п.с.ф.

Данная схема предполагает точное решение задач 1—4. Возможны более или менее детализированные схемы алгоритмов § 7—9, связанные с непосредственным учетом точностей ε приближенного решения задач 1—4. Однако при теоретическом анализе этих схем и их практическом применении возникает ряд проблем. Прежде всего, некоторые из оценок точностей ε на практике оказываются неизвестными или получение таких оценок затруднительно (например, для некоторых функционалов J_δ, Ω трудно дать теоретические оценки точности минимизации сглаживающего функционала). Включать такие оценки в конкретный численный алгоритм поэтому не всегда целесообразно. Далее, сами оценки ε, если они даже известны, зачастую неточны, и необходимо вводить оценки точностей этих оценок. Такой процесс детализации, связанный с введением иерархии точностей, приводит к росту числа параметров в алгоритме и к необходимости связывать эти параметры друг с другом, так же как это, например, делается для итерационных методов [17, 19, 20, 28, 32, 203]. Конечно, идеальной моделью алгоритмов § 4—6 явилось бы описание их работы на уровне машинной арифметики. Однако такой анализ проведен в настоящее время для сравнительно небольшого числа стандартных алгоритмов [44] и представляет собой самостоятельную задачу. Кроме того, модель алгоритмов на уровне машинной арифметики, как и модели с большим числом параметров точности ε, входящих в алгоритм, не всегда целесообразно рассматривать в теоретическом плане.

В некоторых ситуациях (в частности, для многих нелинейных задач) рациональнее вводить параметры ε на уровне программирования алгоритма, если тестовые расчеты показывают целесообразность такого введения.

В данном параграфе будут обсуждены более детально методы решения задач 1—4 при реализации алгоритмов о.п.н., о.п.к., о.п.с.ф.

1. Алгоритмы § 7—9 существенно используют величину λ_δ. В § 3 установлено, что задача вычисления числа λ_δ как задача минимизации по функционалу является устойчивой. Результат теоремы 3.2 обосновывает также возможность вычисления λ_δ с помощью минимизирующих последовательностей. При этом вместо точной величины λ_δ получается ее приближенное значение $\lambda_\eta \equiv \lambda_{\delta\varkappa}$ такое, что $0 \leqslant \lambda_\eta - \lambda_\delta \leqslant \varkappa$. Здесь $\varkappa$ — точность нахождения числа λ_δ, а $\eta \equiv (\delta, \varkappa)$. Анализ результатов, полученных в § 7—9, показывает, что число λ_δ можно заменить на λ_η в определениях функций (7.1), (7.2), (9.1) и соответственно в правилах отбора 7.1, 8.1, 9.1. Величина $\varkappa$ при этом явно не используется. Тогда при выполнении условий теорем 7.2, 7.3, 9.2, 9.3 и при достаточно малых $\|\delta\|$, $\varkappa$ уравнения (7.3), (9.2) имеют

решение $\alpha_\eta \equiv \alpha_{\delta\varkappa} > 0$ (см. доказательства теорем 7.2, 7.3, 9.2, 9.3). Уравнение (8.1) не содержит явно величины λ_η, и для него результаты § 5 о существовании решения α_δ также остаются неизменными. При доказательстве теорем сходимости 7.5, 8.4, 9.6 использовались следующие свойства величины λ_δ: $J^* \leqslant \lambda_\delta$, $\lambda_\delta \to J^*$ при $\Delta \to 0$. Очевидно, теми же свойствами обладает и оценка λ_η: $\lambda_\eta \geqslant \lambda_\delta \geqslant J^*$, $\lambda_\eta \to J^*$ при $\|\eta\| = \|\delta\| + \varkappa \to 0$. Поэтому процедура доказательства теорем сходимости не страдает от замены числа λ_δ на λ_η.

Что касается методов численного нахождения величины λ_η, то этот вопрос тесно связан с проблемой адекватного выбора метода минимизации. Такая же проблема возникает и при рассмотрении п. 2 схемы реализации алгоритмов о.п.н., о.п.к., о.п.с.ф.

2. Касаясь выбора метода минимизации, необходимо отметить, что, несмотря на сложность рассмотрения этой проблемы в общем случае, существует ряд практических и теоретических рекомендаций на этот счет. Из существующей обширной библиографии по этому вопросу отметим обзорные работы [19, 21, 28, 31, 91, 170—173, 215, 219]. Известен ряд численных методов, с успехом применявшихся для решения задач типа задачи о нахождении числа λ_δ или задачи минимизации сглаживающего функционала, в том числе и в нелинейном случае [19, 57, 68, 71, 98, 115, 117, 204—206].

Конкретный выбор адекватного по точности и скорости сходимости численного метода минимизации для задач (4.2), (3.2) зачастую связан с учетом специфики минимизируемого функционала и множества D.

Теоретические исследования в области точности и скорости сходимости конкретных методов минимизации хорошо известны [19, 31, 91, 170—173, 215, 219].

Важную роль в вопросе о способе минимизации сглаживающего функционала играет следующее обстоятельство. Скорость сходимости метода минимизации в принципе зависит от вида используемой в сглаживающем функционале вспомогательной функции f. В случае поиска нормальных решений экстремальной задачи (2.2) на скорость сходимости оказывает влияние также и вид функции g, определяющей регуляризующий функционал $\Omega(z) = g(\|z\|)$. С другой стороны, по крайней мере для варианта алгоритмов о.п.н., о.п.к., о.п.с.ф., рассмотренных в § 13, точность получаемого приближенного решения не зависит от этих функций (см. теорему 13.8). Таким образом, выбирая подходящие функции f, g, можно управлять скоростью сходимости метода минимизации сглаживающего функционала.

В гл. 3, 4, 6 применительно к решению нелинейных операторных уравнений вариационным методом вопрос о выборе базового метода минимизации сглаживающего функционала рассмотрен на конкретной основе опыта решения типичных задач.

Необходимо отметить, что по уровню сложности экстремальная задача (4.2) часто оказывается того же порядка, что и исходная задача (2.2).

3. Обратимся к вопросу о выборе метода решения уравнений с монотонными функциями (7.3), (8.1), (9.2). Основная сложность

связана с возможной разрывностью входящих в них функций. В общем случае это исключает применение таких хорошо известных методов решения уравнений, как метод касательных, метод хорд и др. [21]. Тем не менее такие уравнения могут быть решены, например, методом половинного деления. Этот метод годится и для решения уравнений с разрывными монотонными функциями.

Предположим, что функция $\chi(\alpha)$ определена при $\alpha>0$ и (для определенности) монотонно не убывает там. Будем считать, что уравнение $\chi(\alpha)=0$ имеет решения, принадлежащие известному отрезку $[a, b]$ $(a>0)$, и при этом $\chi(a)<0$, $\chi(b)>0$. Метод половинного деления для решения указанного уравнения связан с построением последовательности вложенных отрезков $[\alpha_n, \beta_n]\subset[a, b]$ $(n=0, 1, \ldots)$, для которых должны выполняться неравенства $\chi(\alpha_n)\leqslant 0$, $\chi(\beta_n)\geqslant 0$. Поскольку случай $\chi(\alpha_n)=0$ или $\chi(\beta_n)=0$ тривиален, то будем считать, что $\chi(\alpha_n)<0$, $\chi(\beta_n)>0$. Переход от n-го к $n+1$-му отрезку осуществляется путем нахождения величины $x_n=(\alpha_n+\beta_n)/2\in(\alpha_n, \beta_n)$ с анализом знака числа $\chi(x_n)$: если $\chi(x_n)<0$, то полагаем $\alpha_{n+1}=x_n$, $\beta_{n+1}=\beta_n$; если $\chi(x_n)>0$, то полагаем $\alpha_{n+1}=\alpha_n$, $\beta_{n+1}=x_n$.

Теорема 1. *Последовательность $\{x_n\}$, построенная по методу половинного деления, сходится к одному из решений уравнения $\chi(\alpha)=0$.*

Доказательство. Для последовательностей $\{\alpha_n\}$, $\{\beta_n\}$, $\{x_n\}$, определенных в методе половинного деления, справедливы следующие неравенства:

$$\alpha_n\leqslant\alpha_{n+1}, \quad \beta_{n+1}\leqslant\beta_n; \quad \beta_{n+1}-\alpha_{n+1}\leqslant(\beta_n-\alpha_n)/2;$$
$$\chi(\alpha_n)<0, \ \chi(\beta_n)>0.$$

Из этих неравенств по обычной схеме обоснования метода половинного деления получаются сходимости $\alpha_n\uparrow c-0$, $\beta_n\downarrow c+0$ при $n\to\infty$. Тогда из неубывания функции $\chi(\alpha)$ получим

$$\chi(\alpha_n)\to\chi(c-0)\leqslant 0, \quad \chi(\beta_n)\to\chi(c+0)\geqslant 0.$$

Неравенства $\chi(c-0)\leqslant 0$, $\chi(c+0)\geqslant 0$ доказывают, что число c есть или обычное решение уравнения, или точка скачка функции $\chi(\alpha)$ через нулевое значение. Таким образом, число c есть решение рассматриваемого уравнения (см. определение 6.1). Сходимость $x_n\to c$ при $n\to\infty$ вытекает из неравенства $\alpha_n\leqslant x_n\leqslant\beta_n$.

Для эффективного решения уравнения типа $\chi(\alpha)=0$ необходима оценка для границ отрезка $[a, b]$ локализации его решений. Во многих задачах такие оценки можно найти [41, 58, 82, 226].

4. Рассмотрим вопросы отбора экстремали при решении уравнения $\chi(\alpha)=0$, где $\chi(\alpha)$ — одна из функций, входящих в уравнение (7.3), (8.1) или (9.2). Если бы решение уравнения α_δ было известно точно, то в принципе можно было бы находить приближенное решение z^{α_δ} задачи (2.4) в соответствии с правилами отбора 7.1, 8.1 или 9.1. С вычислительной точки зрения, однако, это потребовало бы в общем случае многократного решения экстремальной задачи (4.2) минимизации сглаживающего функционала при $\alpha=\alpha_\delta$ с целью нахождения нужной экстремали. Между тем само число α_δ находится приближенно,

например, методом половинного деления. Поэтому правило отбора должно учитывать факт неточного нахождения параметра α_δ. В связи с этим рассмотрим модификации правил отбора, учитывающие приближенность определения α_δ и использующие однократное решение задачи (4.2) при каждом рассматриваемом α. Эти модификации специально предназначены для использования в вычислительном процессе [135].

Будем считать, что имеется последовательность приближений $\alpha_n \equiv \alpha_n(\delta)$, $\beta_n \equiv \beta_n(\delta)$ для решения $\alpha_\delta > 0$ уравнения $\chi(\alpha) = 0$, причем выполнены условия $0 < \alpha_n < \alpha_\delta < \beta_n$, $\chi(\alpha_n)\,\chi(\beta_n) < 0$ для всякого номера n и $\alpha_n \to \alpha_\delta - 0$, $\beta_n \to \alpha_\delta + 0$ при $n \to \infty$. Такими свойствами обладают, например, последовательности α_n, β_n, построенные в методе половинного деления. Фиксируем параметр алгоритма $q > 1$. Тогда существует такой $n(\delta)$, для которого выполнено неравенство

$$\beta_{n(\delta)}/q < \alpha_\delta \qquad (\alpha_{n(\delta)}q > \alpha_\delta). \tag{1}$$

При выполнении этого условия введем вспомогательные значения параметра регуляризации $a(\delta) \equiv \beta_{n(\delta)}/q$, $b(\delta) \equiv \beta_{n(\delta)}$ $(a(\delta) \equiv \alpha_{n(\delta)}$, $b(\delta) \equiv \alpha_{n(\delta)}q)$. По аналогии с используемыми в правилах отбора 7.1, 8.1, 9.1 величинами введем числа $b_1(\delta) = b(\delta)/q = a(\delta)$, $b_2(\delta) \equiv b(\delta)q$ $(a_1(\delta) \equiv a(\delta)/q$, $a_2(\delta) \equiv a(\delta)q = b(\delta))$, а также любые соответствующие этим значениям параметра регуляризации экстремали $z^{b_{1,2}}$ $(z^{a_{1,2}})$ задачи (4.2). Вследствие выбора величин $a(\delta)$, $b(\delta)$ и на основании (1) и монотонности функции $\chi(\alpha)$ можно утверждать, что выполнены неравенства *): $a(\delta) < \alpha_\delta < b(\delta)$, $\chi\,[a(\delta)]\,\chi\,[b(\delta)] < 0$.

Сформулируем два аналога правила отбора 7.1 для алгоритма о.п.н.

Правило отбора 1. Если выполнено неравенство

$$\begin{gathered} I_\delta(z^{b_2}) \geqslant C\Pi_\delta(z^{b_1}) - (C-1)f(\lambda_\delta) \\ (I_\delta(z^{a_2}) \geqslant C\Pi_\delta(z^{a_1}) - (C-1)f(\lambda_\delta)), \end{gathered} \tag{2}$$

то в качестве приближения к решению $\alpha_\delta > 0$ уравнения (7.3) принимаем величину $\alpha(\delta) \equiv a(\delta)$ и в качестве приближения к множеству Ω-оптимальных решений задачи (2.2) выбираем $z_\delta \equiv z^{a(\delta)}$. При этом из (7.3), леммы 6.1 и следствия 6.5 следует неравенство

$$I_\delta(z_\delta) \leqslant \Pi_\delta(z_\delta). \tag{3}$$

Если же выполнено неравенство

$$\begin{gathered} I_\delta(z^{b_2}) \leqslant C\Pi_\delta(z^{b_1}) - (C-1)f(\lambda_\delta) \\ (I_\delta(z^{a_2}) \leqslant C\Pi_\delta(z^{a_1}) - (C-1)f(\lambda_\delta)), \end{gathered} \tag{4}$$

то полагаем $\alpha(\delta) \equiv b(\delta)$, $z_\delta \equiv z^{b(\delta)}$. При этом справедливо неравенство

$$I_\delta(z_\delta) \geqslant \Pi_\delta(z_\delta). \tag{5}$$

---

*) Если $\chi\,[a(\delta)]\,\chi\,[b(\delta)] = 0$, то одно из чисел $a(\delta)$, $b(\delta)$ — обычное решение уравнения $\chi(\alpha) = 0$. Его и соответствующую ему экстремаль задачи (4.2) следует взять как α_δ и z_δ.

Правило отбора 1 оперирует с двумя значениями параметра регуляризации ($b_1=a(\delta)$, $b_2=b(\delta)q$ или $a_1=a(\delta)/q$, $a_2=b(\delta)$) и с двумя произвольными, соответствующими им экстремалями задачи (4.2) ($z^{b_1}=z^{a(\delta)}$, z^{b_2} или z^{a_1}, $z^{a_2}=z^{b(\delta)}$). При этом одно из значений параметра регуляризации принимается за приближенное решение уравнения (7.3), а соответствующая экстремаль — за приближенное решение задачи (2.4). Неравенства (3), (5) не требуют проверки и автоматически следуют из условий выбора экстремалей z_δ. Справедлива следующая теорема сходимости приближений.

Теорема 2. *Пусть выполнены предположения* 1—3 *из* § 2, *условия теоремы* 7.3 *и последовательность* $\{\delta_m\}_{m=1}^{\infty}$ *такова, что* $\|\delta_m\|\to 0$ *при* $m\to\infty$. *Тогда соответствующая последовательность приближений* $\{z_m\}$: $z_m\equiv z_{\delta_m}$, *построенная по правилу отбора* 1, *обладает свойствами*

$$z_m\xrightarrow{\tau}\bar{Z},\quad \Omega(z_m)\to\bar{\Omega},\quad m\to\infty.$$

Доказательство обрисуем схематически. Соотношения (2)—(5) аналогичны неравенствам (7.7)—(7.10) из правила 7.1 и соответствуют условиям (5.13)—(5.16). Нетрудно проверить, пользуясь леммой 6.1 и неравенствами (5.1), (5.2), что для величин $\alpha(\delta)>0$, z_δ, z^{b_1}, z^{b_2} ($\alpha(\delta)>0$, z_δ, z^{a_1}, z^{a_2}) выполнены и остальные условия а)—д) из § 5, в которых положено $\bar{\Omega}_\delta\equiv\bar{\Omega}$. Таким образом, для z_δ оказываются выполненными условия лемм 5.7—5.10, из которых следует, что для семейства z_δ справедливы соотношения (5.31), (5.32). Тогда по лемме 5.1 и следствию 5.1 для z_δ справедлив и аналог теоремы сходимости 7.5, представляющий собой теорему 2.

Способ выбора приближенного решения, близкий к правилу отбора 1, изучался в работе [29].

По такой же методике можно сформулировать и аналоги правила отбора 8.1 для алгоритма о.п.к.

Правило отбора 2. Если выполнено (2), то в качестве приближения для решения $\alpha_\delta>0$ уравнения (8.1) примем $\alpha(\delta)=a(\delta)$, а в качестве приближенного решения задачи (2.4) — элемент $z_\delta=z^{a(\delta)}$. При этом из леммы 6.1 и следствия 6.5 вытекает неравенство

$$\Omega(z_\delta)\geqslant\omega_\delta. \tag{6}$$

Если же выполнено (4), то положим $\alpha(\delta)=b(\delta)$, $z_\delta=z^{b(\delta)}$, и, таким образом, будет верно неравенство

$$\Omega(z_\delta)\leqslant\omega_\delta. \tag{7}$$

Для величин $\alpha(\delta)$, z_δ, z^{b_1}, z^{b_2} ($\alpha(\delta)$, z_δ, z^{a_1}, z^{a_2}), входящих в это правило, будут выполнены условия а)—в), д) из § 5 с $\bar{\Omega}_\delta=\omega_\delta>\bar{\Omega}$, а также условия леммы 5.11. Тогда по этой лемме для семейства z_δ выполнены соотношения (5.31), (5.32), и из следствия 5.1 вытекает

Теорема 3. *Пусть выполнены условия теоремы* 8.4 *и предположения* 1—3 *из* § 2. *Тогда для любой последовательности* $\{\delta_m\}$: $\delta_m\to 0$ *при* $m\to 0$ *соответствующая последовательность* $\{z_m\}$: $z_m\equiv z_{\delta_m}$ *приближенных решений задачи* (2.4), *полученная по правилу отбора* 2, *обладает свойствами сходимости*: $z_m\xrightarrow{\tau}\bar{Z}$, $\Omega(z_m)\to\bar{\Omega}$ *при* $m\to\infty$.

Сформулируем, наконец, аналог правила 9.1 для алгоритма о.п.с.ф. Пусть $\Omega^* \geqslant 0$.

Правило отбора 3. В качестве приближенного значения решения $\alpha_\delta > 0$ уравнения (9.2) примем число $\alpha(\delta) = b(\delta)$, а в качестве приближенного решения задачи (2.4) — элемент $z_\delta = z^{b(\delta)}$. При этом будет выполнено (по теореме 9.1 и следствию 6.5) неравенство $E_\delta(z_\delta) \geqslant 0$.

Это правило оперирует с одним значением $\alpha(\delta)$ параметра регуляризации и с одной соответствующей ему экстремалью z_δ задачи (4.2).

Из (1) ясно, что для $a(\delta)$ выполнено неравенство $a(\delta) = \beta_{n(\delta)}/q < \alpha_\delta$. Поскольку по теореме 9.4 уравнение (9.2) имеет единственное решение α_δ и по теореме 9.5 имеем $\alpha_\delta \to 0$ при $\Delta \to 0$, то и $a(\delta) \to 0$. Но тогда и $b(\delta) \equiv a(\delta)q$ также сходится к нулю при $\Delta \to 0$. Используя это и рассматривая схему доказательства теоремы сходимости 9.6, можно понять, что она остается справедливой и при замене α_δ на $\alpha(\delta)$ и z^{α_δ} на z_δ. Тогда можно установить следующую теорему.

Теорема 4. *Пусть выполнены условия теоремы* 9.6 *и произвольная последовательность* $\{\delta_m\}$ *сходится к нулю при* $m \to \infty$. *Тогда для соответствующей последовательности* $z_m = z_{\delta_m}$, *выбранной по правилу отбора* 3, *выполнены соотношения* $z_m \xrightarrow{\tau} \bar{Z}$, $\Omega(z_m) \to \bar{\Omega}$ *при* $m \to \infty$.

Таким образом, предлагаемые аналоги правил отбора для случая приближенного нахождения α_δ также обеспечивают τ-сходимость приближений к множеству Ω-оптимальных решений задачи (2.2), т.е. порождают регуляризирующий алгоритм.

Для применения правил отбора 1—3 достаточно сделать $n(\delta)$ итераций по нахождению корня α_δ соответствующего уравнения используемого для выбора параметра регуляризации, с тем чтобы обеспечить выполнение условия (1). Условие (1) легко проверяется в процессе решения уравнения $\chi(\alpha) = 0$, ибо вследствие монотонности функции $\chi(\alpha)$ оно следует из выполнения неравенства $\chi[\beta_{n(\delta)}/q]\,\chi[\beta_{n(\delta)}] < 0$ ($\chi[\alpha_{n(\delta)}]\,\chi[\alpha_{n(\delta)}q] < 0$). Необходимо заметить, что мы рассматриваем наиболее сложный случай, когда величины $\chi[\beta_{n(\delta)}]$, $\chi[\beta_{n(\delta)}/q]$ ($\chi[\alpha_{n(\delta)}]$, $\chi[\alpha_{n(\delta)}q]$) не обращаются в нуль. Случай их обращения в нуль прост не только по той причине, что точно находится обычный корень α_δ, но и вследствие того, что нет необходимости применять правила отбора. В этом случае можно использовать в качестве приближения к $\bar{Z}$ любую экстремаль z^{α_δ} (см. § 7—9).

Можно оценить число итераций $n(\delta)$ для метода половинного деления, достаточное для выполнения условия (1).

Теорема 5. *Пусть известен начальный отрезок* $[a_0, b_0]$ $(a_0 > 0)$ *локализации корня* $\alpha_\delta > 0$ *уравнения* $\chi(\alpha) = 0$. *Тогда при*

$$n(\delta) > \log_2\{(b_0 - a_0)/[a_0(q-1)]\} \tag{8}$$

условие (1) *выполнено.*

Действительно, из условия (8) легко получить, что

$$(b_0-a_0)/2^{n(\delta)}\leqslant a_0(q-1)<\alpha_{n(\delta)}(q-1),$$

откуда получается условие (1):

$$\beta_{n(\delta)}/q\leqslant[\alpha_{n(\delta)}+(b_0-a_0)/2^{n(\delta)}]/q\leqslant\alpha_{n(\delta)}<\alpha_\delta.$$

В заключение главы отметим, что вопросы конечномерной аппроксимации экстремальных задач, влияние этой аппроксимации на вид алгоритмов о.п.н., о.п.к., о.п.с.ф. и связанные с этим теоретические проблемы будут рассмотрены в гл. 3. Там же изучаются детали численной реализации конечномерных алгоритмов с учетом сказанного в данном параграфе и приводятся некоторые результаты модельных расчетов.

ГЛАВА 2

ВАРИАЦИОННЫЕ АЛГОРИТМЫ РЕШЕНИЯ НЕЛИНЕЙНЫХ ОПЕРАТОРНЫХ УРАВНЕНИЙ

§ 1. Постановка задачи и схема ее решения

Во многих случаях прикладные задачи могут быть сформулированы на математическом языке не в вариационной форме, а в виде операторного уравнения. При этом в отличие от задач, которые изучались в гл. 1, возникает определенная специфика, которую можно использовать для упрощения алгоритмов. Для задач в форме операторных уравнений несколько видоизменяются и некоторые понятия, введенные в гл. 1 для некорректных экстремальных задач.

Предположим, что (Z, τ) — топологическое пространство, а U — метрическое пространство, наделенной метрикой ρ. Пусть оператор A (в общем случае нелинейный) определен на некотором множестве $D \subset Z$ и действует из D в U. Рассмотрим задачу решения операторного уравнения на множестве D: при некотором $u \in U$ найти элемент $z \in D$ такой, что

$$Az = u. \tag{1}$$

Как правило, точные данные (A, u) задачи (1) неизвестны, а вместо них заданы приближенные данные (A_h, u_σ). Здесь A_h — приближенный оператор, действующий из D в U и принадлежащий некоторому классу $\mathscr{A}$ допустимых приближенных операторов таких, что для любого $A_h \in \mathscr{A}$

$$\lim_{h \to 0} \rho(A_h z, Az) = 0 \quad \forall z \in D.$$

Индекс h определяет «погрешность» приближенного оператора. Приближенная правая часть $u_\sigma \in U$ такова, что

$$\lim_{\sigma \to 0} \rho(u_\sigma, u) = 0.$$

Число σ связано с «погрешностью» задания правой части уравнения (1).

Определение 1. Задача (1) называется *корректно поставленной* если:

1) она имеет решение, принадлежащее множеству D;
2) это решение единственно;

3) решение обладает свойством устойчивости по отношению к возмущениям данных (A, u): для любой последовательности $\delta_n \equiv (h_n, \sigma_n)$, сходящейся к нулю при $n \to \infty$, и для соответствующей последовательности $\{(A_{h_n}, u_{\sigma_n})\}$: $A_{h_n} \in \mathscr{A}$, $u_{\sigma_n} \in U$, допустимых приближенных данных уравнения

$$A_{h_n} z = u_{\sigma_n}$$

разрешимы на множестве D при $n \geqslant n_0 = \text{const}$ и множества Z_n их решений сходятся в топологии τ к решению задачи (1) при $n \to \infty$.

Хорошо известно [88, 110, 202], что операторные уравнения типа (1) могут представлять собой некорректно поставленную задачу. Это связано прежде всего с отсутствием обычного решения уравнения (1) на множестве D. В связи с этим в [84, 85] было введено понятие квазирешения операторного уравнения на множестве D.

Определение 2. *Квазирешением уравнения* (1) *на множестве* D называется всякий элемент $z^* \in D$, для которого

$$\rho(Az^*, u) = \inf\{\rho(Az, u):\ z \in D\} \equiv \mu_0. \tag{2}$$

Число μ_0 называется *мерой несовместности* операторного уравнения (1) на множестве D.

В частном случае, когда $D = Z$, решения экстремальной задачи (2), если они существуют, называются *псевдорешениями* уравнения (1). Квазирешения, для которых оказывается выполненным равенство $\mu_0 = 0$, совпадают с обычными решениями уравнения (1). Задача о нахождении квазирешений на множестве D как вариационная задача может оказаться некорректно поставленной. В частности, квазирешения существуют не для всяких задач типа (1). В работах [78, 87, 88] исследованы условия корректности постановки задач о поиске квазирешений.

В дальнейшем будем считать, что для правой части $u = \bar{u} \in U$ уравнения (1) множество $Z^* \subset D$ квазирешений задачи (1) непусто, хотя, быть может, состоит и более чем из одного элемента.

Предположим, что на множестве D задан ограниченный снизу функционал $\Omega(z)$: для любого $z \in D$

$$\Omega(z) \geqslant \Omega^* \equiv \inf\{\Omega(z): z \in D\} > -\infty.$$

Поставим задачу о поиске Ω-оптимальных квазирешений задачи (1): найти элементы $\bar{z} \in Z^*$ такие, для которых

$$\Omega(\bar{z}) = \inf\{\Omega(z): z \in Z^*\} \equiv \bar{\Omega}. \tag{3}$$

Обозначим множество таких решений задачи (3) как $\bar{Z}$. Ниже будут указаны условия, при которых $\bar{Z}$ непусто. Если $D = Z$, то о множестве $\bar{Z}$ говорят как о множестве Ω-оптимальных псевдорешений задачи (1).

Уточним теперь класс приближенных данных (A_h, u_σ) задачи (1), которые будут использоваться в дальнейшем. Будем предполагать, что кроме элемента $u_\sigma \in U$ известно число $\sigma \geqslant 0$ — оценка погрешности правой части уравнения (1), причем выполнено неравенство

$$\rho(\bar{u}, u_\sigma) \leqslant \sigma. \tag{4}$$

Рассмотрим также специальное множество $\mathscr{A}$ допустимых приближенных операторов A_h, действующих из D в U: для всякого $A_h \in \mathscr{A}$ должно выполняться условие аппроксимации

$$\rho(Az, A_h z) \leqslant \psi(h, \Omega(z)) \quad \forall z \in D. \tag{5}$$

Здесь мера аппроксимации $\psi(h, \Omega)$ известна, а числовая характеристика $h \geqslant 0$ «близости» операторов A и A_h также считается заданной. В целом «близость» приближенных данных (A_h, u_σ) к точным $(A, \bar{u})$ характеризуется вектором $\delta = (h, \sigma) \in \mathbf{R}^2_+$. При этом $(A_h, u_\sigma) = (A, \bar{u})$, если $\delta = (0, 0)$.

Основная задача, изучаемая в данной главе, заключается в нахождении по совокупности $(A_h, u_\sigma, \psi, h, \sigma)$ приближенных данных задачи (1) элемента $z_\delta \equiv z_\delta(A_h, u_\sigma, \psi, h, \sigma) \in D$, который τ-секвенциально сходился бы при $\delta \to 0$ к множеству $\bar{Z}$ Ω-оптимальных квазирешений задачи (1).

Будет также рассматриваться и задача об устойчивой оценке меры несовместности уравнения (1) на множестве D: по данным $(A_h, u_\sigma, \psi, h, \sigma)$ найти такое число λ_δ, которое сходится к μ_0 при $\delta = (h, \sigma) \to 0$.

Основная задача и задача об оценке меры несовместности могут быть решены на основе вариационного подхода к решению операторных уравнений с использованием схем, изложенных в § 1.3, 1.4. Введем функционалы

$$J(z) \equiv \rho(Az, \bar{u}), \quad J_\delta(z) \equiv \rho(A_h z, u_\sigma), \tag{6}$$

определенные и ограниченные снизу на D. Из (4), (5) легко получить неравенство

$$|J(z) - J_\delta(z)| \leqslant \rho(Az, A_h z) + \rho(\bar{u}, u_\sigma) \leqslant \psi(h, \Omega(z)) + \sigma \equiv \Psi(\delta, \Omega(z)), \tag{7}$$

представляющее собой вариант условия аппроксимации (1.2.3). Таким образом, основную задачу данной главы можно считать частным случаем общей задачи, поставленной в § 1.2: по приближенным данным (J_δ, Ψ, δ) построить элемент $z_\delta \in D$, τ-секвенциально сходящийся к множеству $\bar{Z}$ решений задачи (3) при $\delta \to 0$. Используем для решения этой задачи схему, рассмотренную в § 1.4. В ее основе лежит использование сглаживающего функционала

$$M^\alpha[z] = \alpha\Omega(z) + f[\rho(A_h z, u_\sigma)], \quad z \in D, \quad \alpha > 0, \tag{8}$$

и экстремальной задачи: при заданном $\alpha > 0$ найти элементы $z^\alpha \equiv z^\alpha_\delta \in D$, такие, что

$$M^\alpha[z^\alpha] = \inf\{M^\alpha[z] : z \in D\}.$$

Здесь $f(x)$ — вспомогательная функция. Пусть $Z^\alpha \equiv Z^\alpha_\delta$ — множество решений задачи (8) при данном $\alpha > 0$, предполагаемое непустым. Схема построения приближений к множеству Ω-оптимальных квазирешений задачи (1) включает специальный выбор параметра регуляризации α_δ, фиксацию соответствующего множества Z^{α_δ} решений задачи (9) и выбор из него по специальному правилу отбора элемента $z^{\alpha_\delta} \equiv z^{\alpha_\delta}_\delta \in D$, принимаемого в качестве приближенного решения.

Касаясь задачи нахождения устойчивой оценки меры несовместности, заметим, что и она является частным случаем задачи устойчивого определения величины точной нижней грани функционала (см. § 1.3). Она может быть решена на основе подхода, предложенного в § 1.3. При этом в качестве приближения к величине μ_0, отвечающего данным $(A_h,\ u_\sigma,\ \psi,\ h,\ \sigma)$, принимается число

$$\lambda_\delta=\inf\{J_\delta(z)+\Psi(\delta,\ \Omega(z)):\ z\in D\}=\inf\{\rho(A_hz,\ u_\sigma)+\psi(h,\ \Omega(z))+\sigma:\ z\in D\}. \tag{10}$$

Число λ_δ обычно называют *обобщенной мерой несовместности* задачи (1). Обобщенная мера несовместности была введена в работах [124, 125], а затем изучалась также для линейного случая в [39, 104—106].

Сформулируем основные предположения, используемые в данной главе.

1. Считаем, что в U задана топология t, в общем случае отличающаяся от топологии, порождаемой метрикой ρ. Метрика ρ по предположению обладает свойством: для любого фиксированного $u_0\in U$ функционал $\rho(u,\ u_0)$ t-секвенциально полунепрерывен снизу в U по аргументу $u\in U$.

2. Операторы A, A_h считаются τ-секвенциально непрерывными из D в $(U,\ t)$: для любой последовательности $\{z_n\}\subset D$ такой, что $z_n\xrightarrow{\tau}z_0\in D$, имеет место сходимость $Az_n\xrightarrow{t}Az_0$ при $n\to\infty$.

3. Функционал Ω и функция $\psi(h,\ \Omega)$, определенная при $h\geqslant 0$ и $\Omega\geqslant\Omega^*$, отвечают основным предположениям 1—3 из § 1.2.

4. Функция $f(x)$ является сильно возрастающей некоторого порядка $m\geqslant 1$.

Выполнение предположений 1—3 гарантирует выполнение предположений 1, 2 из § 1.2 для функционалов J, J_δ, Ω, а также предположения 3 из § 1.2 для меры аппроксимации $\Psi(\delta,\ \Omega)$, введенной в формуле (7). Таким образом, для задачи приближенного отыскания Ω-оптимальных квазирешений уравнения (1) выполнен комплекс предположений, обеспечивающий применение для этой задачи методов гл. 1.

Из этих предположений непосредственно вытекают следующие факты:

а) для числа λ_δ справедлива теорема 1.3.1: $\lambda_\delta\geqslant\mu_0$, $\lambda_\delta\to\mu_0$ при $\delta\to 0$, т. е. величина (10) дает устойчивую оценку для меры несовместности;

б) множество $\bar{Z}$ Ω-оптимальных квазирешений задачи (1) на множестве D непусто (см. теорему 1.4.1);

в) множество Z^α решений задачи (9) непусто для каждого $\alpha>0$ (см. теорему 1.4.2).

Справедливы также и другие вспомогательные утверждения, касающиеся функционалов (6) и Ω, которые в общем случае доказаны в § 1.3—1.6.

Из предположений 1—4 следует также возможность применения алгоритмов из § 7—9 для отыскания приближений к множеству Ω-оптимальных квазирешений. Конкретизация этих алгоритмов для задачи (1), (3) будет дана в § 2.

§ 2. Алгоритмы о.п.н., о.п.к., о.п.с.ф. для операторных уравнений

Используя экстремали z^α задачи (1.9) при фиксированном $\delta=(h, \sigma)$ ($\delta\neq(0, 0)$), определяем по аналогии с § 1.6—1.9 вспомогательные функции

$$\gamma(\alpha)\equiv\gamma_\delta(\alpha)=\Omega(z^\alpha),\quad \beta(\alpha)\equiv f[\rho(A_h z^\alpha, u_\sigma)]\equiv I_\delta(z^\alpha),$$
$$\varphi(\alpha)\equiv M^\alpha[z^\alpha],\ \pi(\alpha)\equiv f[\psi(h, \Omega(z^\alpha))+\sigma+\lambda_\delta]\equiv\Pi_\delta(z^\alpha),$$
$$\rho(\alpha)\equiv\beta(\alpha)-\pi(\alpha)\equiv P_\delta(z^\alpha)=I_\delta(z^\alpha)-\Pi_\delta(z^\alpha),$$
$$\varepsilon(\alpha)\equiv \mathrm{E}(z^\alpha)=\varphi(\alpha)-f\{\lambda_\delta+[\psi(h, \Omega(z^\alpha))+\sigma]^p\},$$
$$p=\mathrm{const},\ 0<p<1,\ \alpha>0\quad \forall z^\alpha\in Z^\alpha.$$

Функция $\beta(\alpha)$ в случае, когда $f(x)=x$, называется *невязкой*, а функция $\rho(\alpha)$ — *обобщенной невязкой* для задачи (1.1).

На основании сказанного в § 1 можно утверждать, что все эти функции обладают теми же свойствами, что и их аналоги, которые изучались в § 1.6—1.9. В частности, на основании лемм 1.6.4, 1.6.6 можно написать равенства

$$\lim_{\alpha\to+0}\rho(\alpha)=f(\mu_\delta)-f[\psi(h, \Omega_\delta)+\sigma+\lambda_\delta]\equiv\rho_0,$$
$$\lim_{\alpha\to+\infty}\rho(\alpha)=\nu_\delta-f[\psi(h, \Omega^*)+\sigma+\lambda_\delta]\equiv\rho_\infty,$$
$$\lim_{\alpha\to+0}\varepsilon(\alpha)=f(\mu_\delta)-f\{[\psi(h, \Omega_\delta)+\sigma]^p+\lambda_\delta\}\equiv\varepsilon_0,$$
$$\lim_{\alpha\to+\infty}\varepsilon(\alpha)=\nu_\delta-f\{[\psi(h, \Omega^*)+\sigma]^p+\lambda_\delta\}\equiv\varepsilon_\infty.$$

Здесь введены обозначения

$$\Omega_\delta\equiv\lim_{\alpha\to+0}\Omega(z^\alpha),\quad \mu_\delta=\inf\{\rho(A_h z, u_\sigma)\colon z\in D\},$$
$$\nu_\delta\equiv\inf\{f[\rho(A_h z, u_\sigma)\colon z\in Z_0\},\quad Z_0\equiv\{z\in D\colon \Omega(z)=\Omega^*\}.$$

Сформулируем *обобщенный принцип невязки* для решения уравнения (1.1).

1. Параметр регуляризации α_δ выбирается как решение уравнения с монотонно неубывающей функцией

$$\rho(\alpha)=0. \tag{2}$$

2. Приближенное решение z^{α_δ} выбирается из множества Z^{α_δ} экстремалей задачи (1.9) при $\alpha=\alpha_\delta$ по следующему в а р и а н т у 2.1 п р а в и л а о т б о р а 1.7.1.

Пусть $q>1$, $C>1$ — заданные константы, $\alpha_1\equiv\alpha_\delta/q$, $\alpha_2\equiv\alpha_\delta q$ — вспомогательные значения параметра регуляризации, а z^{α_1}, z^{α_2} — соответствующие им произвольные экстремали задачи (1.9). Если выполнено неравенство

$$I_\delta(z^{\alpha_2})\geqslant C\,\Pi_\delta(z^{\alpha_1})-(C-1)f(\lambda_\delta), \tag{3}$$

то в качестве приближенного решения принимаем любой элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого выполнено условие $P_\delta(z^{\alpha_\delta})\leqslant 0$. В частности, можно принять $z^{\alpha_\delta}=z^{\alpha_\delta}_{-}$ (см. лемму 1.6.5). Если же выполнено неравенство

$$I_\delta(z^{\alpha_2})\leqslant C\Pi_\delta(z^{\alpha_1})-(C-1)f(\lambda_\delta), \tag{4}$$

то за приближенное решение принимаем любой элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого $P_\delta(z^{\alpha_\delta})\geqslant 0$. Например, можно положить $z^{\alpha_\delta}=z^{\alpha_\delta}_{+}$.

Необходимо заметить, что в случае, когда α_δ есть обычное решение уравнения (2), необходимость использования правила отбора исчезает, так как в этом случае выполнено равенство $P_\delta(z^{\alpha_\delta})=\rho(\alpha_\delta)=0$. Тогда в качестве приближенного решения можно принять любую экстремаль $z^{\alpha_\delta}\in Z^{\alpha_\delta}$. То же верно и для алгоритмов о.п.к., о.п.с.ф., формулируемых ниже.

В соответствии с предположениями 1—4 из § 1 и в силу теорем 1.7.1, 1.7.5 справедливы следующие теоремы.

Теорема 1. *Пусть выполнены предположения* 1—4 *из* § 1 *и верны неравенства* $\rho_0<0$, $\rho_\infty>0$. *Тогда уравнение* (2) *имеет решение* $\alpha_\delta>0$. *Неравенства* $\rho_0<0$, $\rho_\infty>0$ *выполняются при достаточно малых* $\|\delta\|$, *если выполнены дополнительные условия: множество* $Z^*\bigcap Z_0$ *пусто, мера аппроксимации* $\psi(h,\Omega)$ *монотонно возрастает по второму аргументу при* $h\neq 0$.

Теорема 2. *Пусть выполнены условия теоремы* 1 *и решение экстремальной задачи* (1.9) *единственно при каждом* $\alpha>0$, *причем для произвольных различных* α_1, $\alpha_2>0$ *реализуется требование* $z^{\alpha_1}\neq z^{\alpha_2}$. *Тогда уравнение* (2) *имеет единственное обычное решение* $\alpha_\delta>0$.

Теорема 3. *При выполнении условий теоремы* 1 *для произвольной последовательности* $\{\delta_n\}$ $(\delta_n\neq(0,0))$, *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательность приближений* $z_n\equiv z^{\alpha_{\delta_n}}$, *полученных по алгоритму о.п.н.,* τ*-сходится к* $\bar{Z}$ *при* $n\to\infty$, *причем* $\Omega(z_n)\to\bar{\Omega}$.

Сформулируем *обобщенный принцип квазирешений* для операторных уравнений (1.1). Он использует кроме данных $\{A_h, u_\sigma, \psi, h, \sigma\}$ оценку ω_δ для неизвестной величины $\bar{\Omega}$.

1. Параметр регуляризации α_δ выбирается как решение уравнения с монотонно невозрастающей функцией

$$\gamma(\alpha)=\omega_\delta. \tag{5}$$

2. Приближенное решение z^{α_δ} выбирается из соответствующего найденному значению α_δ параметра регуляризации множества экстремалей Z^{α_δ} задачи (1.9) по следующему варианту 2.2 правила отбора 1.8.1.

При выполнении условия вида (3) в качестве приближенного решения берется любой элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого справедливо неравенство $\Omega(z^{\alpha_\delta})\geqslant\omega_\delta$; если же выполнено соотношение вида (4), то приближенное решение z^{α_δ} выбирается так, чтобы было справедливым неравенство $\Omega(z^{\alpha_\delta})\leqslant\omega_\delta$.

Из теорем 1.8.1—1.8.4 можно вывести следующие теоремы

Теорема 4. *Пусть выполнены предположения* 1–4 *из* § 1 *и число* ω_δ *удовлетворяет неравенству* $\Omega^*<\omega_\delta<\Omega_\delta$. *Тогда уравнение* (5) *имеет решение* $\alpha_\delta>0$. *Неравенство* $\Omega^*<\Omega_\delta$, *необходимое для положительности числа* α_δ, *выполнено при достаточно малом* $\|\delta\|$, *если удовлетворяются дополнительные условия из теоремы* 1.

Теорема 5. *Предположим, что выполнены условия теоремы 4 и решение задачи* (1.9) *единственно при любом* $\alpha>0$, *причем* $z^{\alpha_1}\neq z^{\alpha}$ *для произвольных различных* α_1, $\alpha_2>0$. *Тогда уравнение* (5) *имеет единственное обычное решение* $\alpha_\delta>0$.

Теорема 6. *Если выполнены условия теоремы* 1 *и, кроме того,* $\omega_\delta\geqslant\bar{\Omega}$, $\omega_\delta\to\bar{\Omega}$ *при* $\|\delta\|\to 0$, *то для любой последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательность* $\{z_n\colon z_n\equiv z^{\alpha_{\delta_n}}\}$, *построенная по алгоритму о.п.к.,* τ-*сходится к* $\bar{Z}$ *и* $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Сформулируем, наконец, *обобщенный принцип сглаживающего функционала* для операторных уравнений. При этом считаем, что $\Omega^*\geqslant 0$.

1. Параметр регуляризации α_δ выбирается как решение уравнения с монотонно не убывающей функцией

$$\varepsilon(\alpha)=0. \tag{6}$$

2. Приближенное решение z^{α_δ} выбирается из множества Z^{α_δ} по следующему варианту 2.3 правила отбора 1.9.1.

За приближенное решение принимается элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого справедливо неравенство $E(z^{\alpha_\delta})\geqslant 0$. В частности, можно положить $z^{\alpha_\delta}=z_{+}^{\alpha_\delta}$.

Из теорем 9.1—9.6 можно получить следующие теоремы

Теорема 7. *Пусть выполнены предположения* 1—4 *из* §1 *и неравенства* $\varepsilon_0<0$, $\varepsilon_\infty>0$. *Тогда уравнение* (6) *имеет единственное решение* $\alpha_\delta>0$. *Неравенства* $\varepsilon_0<0$, $\varepsilon_\infty>0$ *выполняются для достаточно малых* $\|\delta\|$, *если удовлетворяются дополнительные условия из теоремы* 1.

Теорема 8. *Если справедливы предположения* 1—4 *из* §1 *и в дополнение к этому* $Z^*\cap Z=\varnothing$, *то параметр регуляризации* α_δ, *найденный по о.п.с.ф., стремится к нулю при* $\|\delta\|\to 0$.

Теоремы 7, 8 верны и в случае $p=1$.

Теорема 9. *Пусть выполнены предположения* 1—4 *из* §1 *и дополнительные условия из теоремы* 1. *Тогда для любой последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательности приближенных решений* $\{z_n\colon z_n\equiv z^{\alpha_{\delta_n}}\}$, *найденная по о.п.с.ф.,* τ-*сходится к* $\bar{Z}$ *и* $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Как отмечалось в § 1.10, обобщенные принципы невязки, квазирешений и сглаживающего функционала могут формально давать нулевое значение параметра регуляризации α_δ. В этом случае для о.п.к. предлагается в качестве приближенного решения взять любой элемент $z_\delta^0\in D$, минимизирующий на D невязку приближенного уравнения (1):

$$\rho(A_h z_\delta^0,\ u_\sigma)=\inf\{\rho(A_h,\ u_\sigma)\colon z\in D\}, \tag{7}$$

а для о.п.н., о.п.с.ф.— любое решение $z_\delta^0\in Z_0$ экстремальной задачи:

$$\rho(A_h z_\delta^0,\ u_\sigma)=\inf\{\rho(A_h z,\ u_\sigma)\colon z\in Z_0\}. \tag{7'}$$

Возможность такого подхода обосновывается следующей теоремой, вытекающей из теорем 10.1, 10.2.

Теорема 10. *Если выполнены предположения* 1—4 *из* § 1 *и в дополнение к этому мера аппроксимации* $\psi(h,\Omega)$ *возрастает по* *второму аргументу при* $h \neq 0$, *то из равенства* $\alpha_\delta = 0$ *следует разрешимость задач* (7), (7′).

Предположим, что для некоторой последовательности $\{\delta_n\}$, сходящейся к нулю при $n \to \infty$, алгоритмы о.п.н., о.п.к. или о.п.с.ф. дают значение $\alpha_n = \alpha_{\delta_n} = 0$. Тогда из теоремы 10.2 вытекает теорема сходимости приближенных решений $z_n = z^0_{\delta_n}$.

Теорема 11. *При выполнении условий теоремы* 10 *имеет место сходимость* $z_n \xrightarrow{\tau} \bar{Z}$ *и* $\Omega(z_n) \to \bar{\Omega}$, *если* $n \to \infty$.

Сказанное позволяет утверждать, что алгоритмы о.п.н., о.п.к., о.п.с.ф. для операторных уравнений (1) являются регуляризующими и дают решение основной задачи данной главы.

Для операторных уравнений вида (1) удобно приводить примеры, иллюстрирующие не только общую схему вариационной регуляризации, которая дана в § 1.4, 2.1, но и наглядно показывающие свойства вспомогательных функций (1), уравнений (2), (5), (6), а также демонстрирующие «механизм» действия правил отбора для обеспечения сходимости приближений. Рассмотрим некоторые такие примеры.

Пример 1. Зададим нормированное пространство $Z = U = \mathbf{R}$ с нормой $\|z\| = |z|$ ($z \in Z$) и с порождаемой этой нормой топологией $\tau = t$. При этом $\rho(u_1, u_2) = \|u_1 - u_2\| = |u_1 - u_2|$ для любых $u_1, u_2 \in U$. Будем также считать, что $D = Z$. Зафиксируем положительные числа

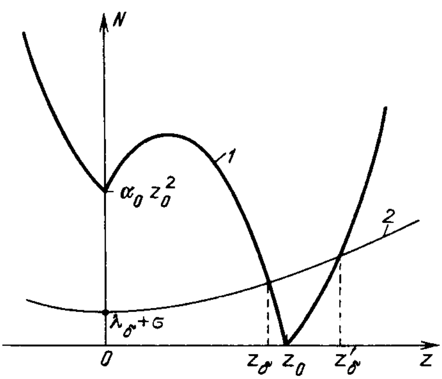

Рис. 5. Иллюстрация к примеру 1: *1* — невязка $N(z) = \|Az - \bar{u}\|^2$; *2* — $\pi(z) = (\lambda_\delta + \sigma + h|z|)^2$

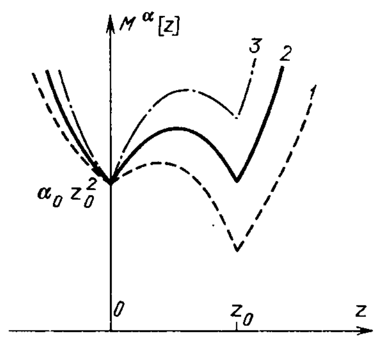

Рис. 6. График сглаживающего функционала: *1* — $0 < \alpha < \alpha_0$; *2* — $\alpha = \alpha_0$; *3* — $\alpha > \alpha_0$

z_0, α_0, $u_0 = \alpha_0^{1/2} z_0$ и введем непрерывный нелинейный оператор $A_0 \colon Z \to U$ по правилу

$$A_0 z = \{[\alpha_0(z_0^2 - z^2) + z(z_0 - z)]^{1/2} + u_0,\ 0 \leqslant z \leqslant z_0;$$
$$[\alpha_0(z^2 - z_0^2)]^{1/2} + u_0,\ z > z_0;\ \alpha_0^{1/2} z_0 + z^2 + u_0,\ z < 0\}.$$

Рассмотрим уравнение (1.1) с точными данными $(A, \bar{u})$: $A = A_0$, $\bar{u} = u_0$. На рис. 5 представлен график невязки $N(z) = \|Az - \bar{u}\|^2$ этого уравнения. Из него ясно, что $\bar{z} = z_0$ представляет собой единственное его решение.

Зададим теперь приближенные данные $(A_h,\ u_\sigma)$, а также величины ψ, h, σ. Пусть h, σ — положительные числа, для которых выполнено неравенство $u_0^2 > 4C(\sigma+h|z_0|)^2$, где $C>1$. Примем $\Omega(z)=\|z\|^2=z^2$, $\psi(h,\Omega(z))=h\|z\|=h|z|$ и будем считать, что $A_h=A_0$, $u_\sigma=u_0$ для любых допустимых h, σ. В качестве вспомогательной функции $f(x)$ будем использовать $f(x)=x^2$. Тогда сглаживающий функционал (1.8) имеет вид

$$M^\alpha[z]=\{(\alpha-\alpha_0)z^2+\alpha_0 z_0^2+z(z_0-z),\ 0\leqslant z\leqslant z_0;$$
$$(\alpha+\alpha_0)z^2-\alpha_0 z_0^2,\ z>z_0;\ \alpha z^2+(\alpha_0^{1/2}z_0+z^2)^2,\ z<0\}.$$

Графики функций $M^\alpha[z]$ для характерных значений параметра α приведены на рис. 6. Понятно, что элементы вида $z^\alpha=\{z_0$ при $0<\alpha<\alpha_0$; $0,\ z_0$ при $\alpha=\alpha_0$; 0 при $\alpha>\alpha_0\}$ являются решением задачи (1.9) минимизации сглаживающего функционала. Это позволяет вычислить вспомогательные функции (1), графики которых представлены на рис. 7. В частности, обобщенная невязка $\rho(\alpha)$ задается равенством

$$\rho(\alpha)=\{\rho_0,\ 0<\alpha\leqslant\alpha_0;\ \rho_\infty,\ \alpha\geqslant\alpha_0\},$$

где $\rho_0=-(\lambda_\delta+\sigma+h|z_0|)^2$, $\rho_\infty=|u_0|^2-(\sigma+\lambda_\delta)^2$. Число λ_δ в данной задаче оценивается следующим образом: $0<\lambda_\delta=\inf\{\|A_0z-u_0\|+h\|z\|+\sigma\colon z\in\mathbf{R}\}\leqslant\sigma+h|z_0|$, и поэтому $\rho_\infty>0$.

Характерной особенностью графиков функций (1) является их разрывность в точке $\alpha=\alpha_0$, обусловленная существованием двух

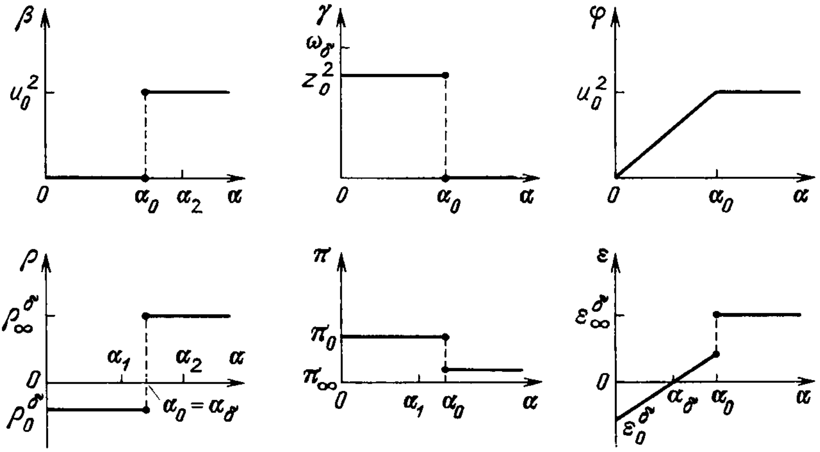

Рис. 7. Графики вспомогательных функций

экстремалей задачи (1.9) при этом значении α: $Z^{\alpha_0}=\{0;\ z_0\}$. Эти экстремали реализуют левые и правые предельные значения функций (1) в точке α_0. Из графика функции $\rho(\alpha)$ легко видеть, что уравнение (2) имеет обобщенное решение $\alpha_\delta=\alpha_0$, и в силу задания функции

$\rho(\alpha)$ это решение не меняется при изменении величин h, $\sigma>0$. Таким образом, значение параметра регуляризации, выбираемое по о.п.н., в общем случае может и не сходиться к нулю при $\delta=(h, \sigma)\to 0$. Для найденного значения $\alpha_\delta=\alpha_0$ задача (1.9) имеет два решения: одна из экстремалей совпадает с точным решением z_0 задачи (1.1), а вторая равна нулю. Поэтому «неудачный» выбор приближенного решения $z^{\alpha_\delta}=0$ привел бы к отсутствию сходимости приближений при $\|\delta\|\to 0$ к точному решению задачи (1.1). Именно это обосновывает необходимость специальных правил отбора экстремалей из множества Z^{α_δ}. Применим правило отбора 2.2.1 алгоритма о.п.н.

Пусть $\alpha_1=\alpha_\delta/q=\alpha_0/q<\alpha_0$, $\alpha_2=\alpha_\delta q=\alpha_0 q>\alpha_0$. Тогда $z^{\alpha_2}=0$, $z^{\alpha_1}=z_0$ и

$$
\begin{aligned}
I_\delta(z^{\alpha_2}) &= \|A_0\cdot 0-u_0\|^2=u_0^2>4C(\sigma+h|z_0|)^2\geqslant \\
&\geqslant C(\lambda_\delta+\sigma+h|z_0|)^2=C(\lambda_\delta+\sigma+\psi(h, \Omega(z^{\alpha_1})))^2= \\
&= C\Pi_\delta(z^{\alpha_1})\geqslant C\Pi_\delta(z^{\alpha_1})-(C-1)f(\lambda_\delta). \qquad (8)
\end{aligned}
$$

Таким образом, выполнено (3), и из множества $Z^{\alpha_\delta}=\{0; z_0\}$ следует выбрать экстремаль, для которой выполняется условие $P_\delta(z^{\alpha_\delta})=\rho(\alpha_\delta)\leqslant 0$. Такой экстремалью является $z^{\alpha_\delta}=z_0$, которая реализует левое предельное значение обобщенной невязки в точке $\alpha=\alpha_\delta$. Итак, алгоритм о.п.н. дает в рассматриваемой задаче приближенное решение $z^{\alpha_\delta}=z_0$, совпадающее с точным решением задачи (1.1).

Выбор параметра регуляризации по обобщенному принципу сглаживающего функционала в рассматриваемом примере приводит к решению уравнения (6). При этом $\varepsilon(\alpha_0-0)=\alpha_0 z_0^2-[\lambda_\delta+(\sigma+h|z_0|)^p]^2$ и при достаточно малых h, σ выполнено условие $\varepsilon(\alpha_0-0)>0$. Это ведет к тому, что уравнение (6) имеет при малых h, σ обычное решение $\alpha_\delta>0$, которому отвечает единственная экстремаль $z^{\alpha_\delta}=z_0$. Ее и следует принять в качестве приближенного решения задачи (1.1), так как $\mathrm{E}(z^{\alpha_\delta})=\varepsilon(\alpha_\delta)=0$.

Если, наконец, применить алгоритм о.п.к., т. е. задаться оценкой сверху ω_δ для величины $\bar{\Omega}=z_0^2$ и рассмотреть уравнение (5), то вследствие соотношения $(\gamma(\alpha)-\omega_\delta)(\alpha-0)=(z_0^2-\omega_\delta)\alpha<0$, верного для любых $\alpha>0$, и по определению решения уравнения (5) оказывается, что $\alpha_\delta=0$. Таким образом, в рассматриваемом примере реализуется случай, когда α_δ (о.п.н.)>0, α_δ (о.п.с.ф.)>0, но α_δ (о.п.к.)$=0$ (ср. лемму 1.10.1). По схеме, развитой в § 1.10, в случае α_δ (о.п.к.)$=0$ принимаем в качестве приближенного решения задачи любой элемент, минимизирующий невязку (7): $\|A_h z-u_\sigma\|=\|A_0 z-u_0\|$, т. е. полагаем $z^{\alpha_\delta}=z_0$.

Пример 1 показывает также, что функции (1) могут и не быть строго монотонными.

Пример 2. Он аналогичен примеру 1 с той лишь разницей, что вместо оператора A_0 берется оператор A_1: $Z\to U$:

$$
A_1 z=\{2u_0,\ z\leqslant 0;\ (u_0^2-\alpha_0 z^2)^{1/2}+u_0,\ 0<z\leqslant z_0;\ (\alpha_0 z^2-u_0^2)^{1/2}+u_0,\ z>z_0\},
$$

где $z_0=u_0\alpha_0^{-1/2}$ — единственное решение уравнения (1.1) с точными данными (A_1, u_0) (рис. 8).

Находя по приближенным данным ($A_h=A_1$, $u_\sigma=u_0, h, \sigma$) экстремали сглаживающего функционала

$$M^\alpha[z]=\{\alpha z^2+u_0^2,\ z\leqslant 0;\ (\alpha-\alpha_0)z^2+u_0^2,\ 0<z\leqslant z_0;\ (\alpha+\alpha_0)z^2-u_0^2,\ z>z_0\}$$

(рис. 9), получаем

$$z^\alpha=z_0,\ 0<\alpha<\alpha_0,\ z^\alpha=0,\ \alpha>\alpha_0,$$

а при $\alpha=\alpha_0$ множество $Z^{\alpha_0}=\{z^{\alpha_0}\}$ состоит из всех точек отрезка $[0, z_0]$. Таким образом, значения функций (1) в их точке разрыва

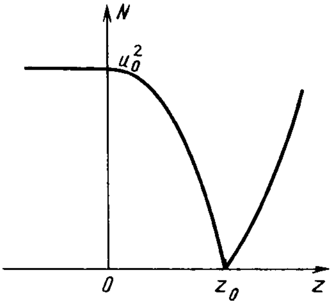

Рис. 8. График функционала невязки

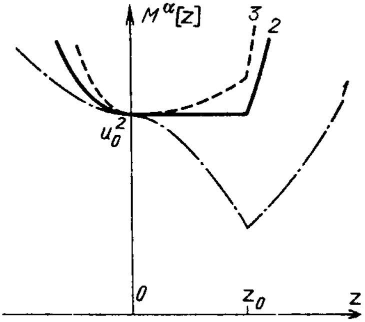

Рис. 9. График сглаживающего функционала: 1 — $0<\alpha<\alpha_0$; 2 — $\alpha=\alpha_0$; 3 — $\alpha>\alpha_0$

$\alpha=\alpha_0$ целиком заполняют отрезок между их предельными значениями слева и справа в этой точке, как это показано на рис. 10.

Используя о.п.н., получаем по аналогии с примером 1: $\alpha_\delta=\alpha_0$, $Z^{\alpha_\delta}=[0, z_0]$, причем $z_-^{\alpha_\delta}=0$, $z_+^{\alpha_\delta}=z_0$. Поскольку в рассматриваемом примере по-прежнему будет выполнено соотношение (8), то отбор экстремали z^{α_δ} должен производиться по первой части правила отбора 2.2.1, т. е. из условия $P_\delta(z^{\alpha_\delta})\leqslant 0$. В нашем конкретном примере оно принимает вид системы неравенств

$$0\leqslant z^{\alpha_\delta}\leqslant z_0,\quad u_0^2-\alpha_0(z^{\alpha_\delta})^2\leqslant(\lambda_\delta+\sigma+hz^{\alpha_\delta})^2,$$

решая которую получаем

$$z(\delta)\equiv\rho_\infty/\left[\sqrt{h^2\sigma^2+(\alpha_0+h^2)\rho_\infty}+\sigma h\right]\leqslant z^{\alpha_\delta}\leqslant z_0.$$

Поэтому в примере 2 в качестве приближенного решения z^{α_δ} можно взять любую точку отрезка $[z(\delta), z_0]$. В частности, элемент $z^{\alpha_\delta}=z(\delta)$ реализует равенство нулю обобщенной невязки.

Следующий пример, касающийся алгоритма о.п.с.ф., показывает, что введение степени p $(0<p<1)$ в правило выбора параметра регуляризации (9.2) не случайно. Если положить $p=1$, то алгоритм о.п.с.ф. в общем случае не будет обеспечивать сходимость приближений к множеству Ω-оптимальных квазирешений уравнения (1.1).

Пример 3. Пусть $Z=D=U=\mathbf{R}^2$, а топологии τ, t порождаются нормой $\|z\|^2=x^2+y^2$ $(z=(x, y)^T\in\mathbf{R}^2)$. Снова считаем, что $\rho(u_1, u_2)=\|u_1-u_2\|$ для всяких u_1, $u_2\in\mathbf{R}^2$ и $f(x)=x^2$. Будем предполагать, что операторы A, A_h: $Z\to U$ линейные и имеют матрицы вида $A=[a_{ij}]$, $A_h=[a_{ij}^h]$ размерности 2×2 с элементами $a_{11}=1$,

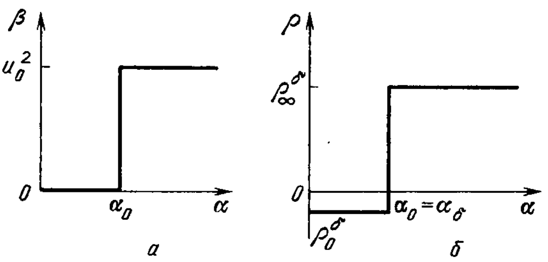

Рис. 10. Графики: *а* — невязки, *б* — обобщенной невязки

$a_{12}=a_{21}=a_{22}=0$; $a_{11}^h=1$, $a_{12}^h=a_{21}^h=0$, $a_{22}^h=h$. Зададим также $\Omega(z)=\|z\|^2$, $\psi(h, \Omega)=h\|z\|$ и положим $\bar{u}=(1, 0)^T$, $u_\sigma=(1, \sigma)^T$. Очевидно, нормальное решение задачи (1.1) для точных данных $(A, \bar{u})$ есть $\bar{z}=(1, 0)^T$. Экстремали сглаживающего функционала

$$M^\alpha[z]=\alpha(x^2+y^2)+(x-1)^2+(hy-\sigma)^2$$

легко вычисляются:

$$z^\alpha=(x^\alpha, y^\alpha), \quad x^\alpha=1/(1+\alpha), \quad y^\alpha=\sigma h/(\alpha+h^2).$$

Напишем условия выбора параметра регуляризации по о. п. с. ф. с $p=1$:

$$\begin{aligned}\varepsilon(\alpha)&=M^\alpha[z^\alpha]-f[\lambda_\delta+\sigma+\psi[h, \|z^\alpha\|]]=\\&=\alpha[(x^\alpha)^2+(y^\alpha)^2]+(x^\alpha-1)^2+(hy^\alpha-\sigma)^2-\left(\lambda_\delta+\sigma+h\sqrt{(x^\alpha)^2+(y^\alpha)^2}\right)^2=\\&=\frac{\alpha}{1+\alpha}+\frac{\alpha\sigma^2}{\alpha+h^2}-\left[\lambda_\delta+\sigma+h\sqrt{\frac{1}{(1+\alpha)^2}+\frac{\sigma^2h^2}{(\alpha+h^2)^2}}\right]^2=0.\end{aligned}$$

Пусть для простоты $\sigma=h$. Тогда из приведенного уравнения получим $\alpha_\delta=\sigma^2A(\sigma)/B(\sigma)$, где $\delta=(\sigma, \sigma)$ и

$$\begin{aligned}A(\sigma)&=\left[(\lambda_\delta/\sigma)+1+\sqrt{1/(1+\alpha_\delta)^2+\sigma^4/(\alpha_\delta+\sigma^2)^2}\right]^2,\\B(\sigma)&=1/(1+\alpha_\delta)+\sigma^2/(\alpha_\delta+\sigma^2).\end{aligned}$$

Из неравенства типа (1.3.4) для рассматриваемой задачи следует, что

$$\lambda_\delta\leqslant f[\rho(A\bar{z}, \bar{u})]+2[\sigma+h\|\bar{z}\|]=4\sigma.$$

Кроме того, по теореме 8 имеем $\alpha_\delta\to 0$ при $\delta\to 0$, так что при $0<\sigma\leqslant\sigma_0$ можно считать выполненным неравенство $\alpha_\delta\leqslant 1$. Поэтому $1\leqslant A(\sigma)\leqslant(5+\sqrt{2})^2$, $1/2\leqslant B(\sigma)\leqslant 2$ для всех σ $(0<\sigma\leqslant\sigma_0)$. Отсюда можно найти оценку для y^{α_δ}:

$$y^{\alpha_\delta}=\frac{\sigma^2}{\alpha_\delta+\sigma^2}=\frac{1}{\alpha_\delta/\sigma^2+1}=\frac{1}{A(\sigma)/B(\sigma)+1}\geqslant\frac{1}{1+\gamma_0}>0, \quad \gamma_0\equiv 2(5+\sqrt{2})^2.$$

Это означает, что приближения $z^{\alpha_\delta}=(x^{\alpha_\delta}, y^{\alpha_\delta})^T$, полученные по о.п.с.ф. с $p=1$, не сходятся при $\delta=(\sigma, \sigma)\to 0$ к нормальному решению $\bar{z}=(1, 0)^T$

В заключение приведем пример, показывающий, что особенности поведения вспомогательных функций для нелинейных задач (разрывность, неоднозначность, не строгая монотонность) могут проявляться и при решении линейных операторных уравнений, если стабилизирующий функционал Ω, удовлетворяет предположениям 1, 2 из § 1.2, но не обладает дополнительным свойством строгой выпуклости (ср. § 1.11).

Пример 4. Пусть снова $Z=D=U=\mathbf{R}$, $\|z\|=|z|$ для любого $z\in\mathbf{R}$. Зададим $A=A_h=E$ для любого h $(0\leqslant h\leqslant h_0)$ и $u_\sigma=\bar{u}=1$ для любого σ $(0\leqslant\sigma\leqslant\sigma_0)$. Предположим также, что $f(x)=x$, а $\Omega(z)=\|z\|$. Легко видеть, что сглаживающий функционал $M^\alpha[z]=\alpha|z|+|z-1|$ в данном примере имеет множество экстремалей

$$Z_\delta^\alpha=\{1\colon 0<\alpha<1;\ \ [0,1]\colon \alpha=1;\ \ 0\colon \alpha>1\}.$$

Соответственно вспомогательные функции (1) принимают вид

$$\varphi(\alpha)=\{\alpha\colon 0<\alpha<1;\ \ 1\colon \alpha\geqslant 1\},$$
$$\gamma(\alpha)=\{1\colon 0<\alpha\leqslant 1;\ \ 0\colon \alpha\geqslant 1\},$$
$$\beta(\alpha)=\{0\colon 0<\alpha\leqslant 1;\ \ 1\colon \alpha\geqslant 1\}.$$

Характерной особенностью этого примера являются не строгая выпуклость функционала Ω, или, говоря другими словами, не строгая нормированность пространства Z.

В § 4 вопрос о влиянии строгой нормированности пространства Z на непрерывность и однозначность вспомогательных функций будет исследован специально.

Таким образом, специфические свойства вспомогательных функций (1), изученные в § 1.6—1.9, 1.11, определяются не только тем, является ли задача (1.1) линейной или нет. Столь же значительны и свойства пространств (Z, τ), U, множества D и функционала Ω.

§ 3. Модификации алгоритмов для случая разрешимых операторных уравнений

Изложенные в § 2 алгоритмы приближенного нахождения Ω-оптимальных квазирешений операторного уравнения (1.1) существенно используют величину λ_δ — устойчивую оценку меры несовместности операторного уравнения. В общем случае исключить число λ_δ из алгоритмов нельзя. Это можно проиллюстрировать следующим примером.

Пример 1. Пусть $Z=D=U=\mathbf{R}$ и топологии τ, t порождены нормой $\|z\|=|z|$ $(z\in\mathbf{R})$. Будем считать, что $\rho(u_1, u_2)=|u_1-u_2|$ для любых u_1, $u_2\in U$. Зададим точный оператор задачи (1.1) $A\colon Z\to U$:

$$Az=\{2\colon z\geqslant 0;\ \ 2-z\colon z<0\}$$

и определим регуляризатор $\Omega(z)=|z+1|$. В качестве точной правой части уравнения (1.1) примем $\bar{u}=1$. Тогда, как легко убедиться, например, из графика невязки $N(z)=\rho(Az, \bar{u})$ (рис. 11) уравнение (1.1) имеет единственное Ω-оптимальное псевдорешение $\bar{z}=0$, причем точная мера несовместности $\mu_0=1$.

Предположим теперь, что вместо точных данных $(A, \bar{u})$ в нашем распоряжении имеются приближенные данные (A_h, u_σ) следующего вида:

$$A_h z=\{2+h-z\colon z<0;\ 2+h(1-\sqrt{h}z)\colon 0\leqslant z\leqslant 1/\sqrt{h};\ 2\colon z>1/\sqrt{h}\},\quad u_\sigma=\bar{u}.$$

При этом считается, что $\sigma=0$, $h>0$, $\delta=(h, 0)$. Если задать меру аппроксимации в виде $\psi(h, \Omega)=h(1+\Omega)$, то легко убедиться, что для данного примера выполнено условие аппроксимации (1.5). Пусть

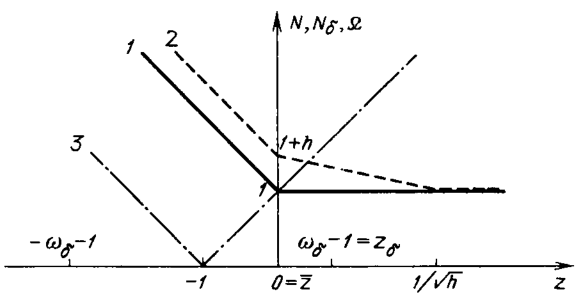

Рис. 11. Графики функционалов: 1 — $N(z)=\|Az-\bar{u}\|^2$; 2 — $N_\delta(z)=\|A_h z-\bar{u}\|^2$; 3 — $\Omega(z)$

$f(x)=x$. Тогда легко найти, что $z^\alpha=\{h^{-1/2}, 0<\alpha\leqslant h^{3/2};\ 0, h^{3/2}\leqslant\alpha\leqslant 1;\ -1, \alpha\geqslant 1\}$. Попробуем использовать алгоритм о. п. н. без величины λ_δ для нахождения $\bar{z}$. Функция обобщенной невязки имеет вид

$$\bar{\rho}(\alpha)=\beta(\alpha)-\psi(h, \Omega(z^\alpha))=$$
$$=\{1-2h-\sqrt{h}\colon 0<\alpha\leqslant h^{3/2}; 1-h\colon h^{3/2}\leqslant\alpha\leqslant 1;\ 2\colon \alpha\geqslant 1\}.$$

Будем считать, что погрешность h достаточно мала и $1-2h-\sqrt{h}>0$. Тогда график обобщенной невязки имеет вид, представленный на рис. 12. Из него ясно, что уравнение $\bar{\rho}(\alpha)=0$ имеет обобщенное решение $\alpha_\delta=0$. В этом случае в качестве приближений к $\bar{z}$ следует принять какое-либо из решений задачи (2.7′), т. е. $z_\delta^0=-1$. Это исключает сходимость z_δ^0 к $\bar{z}$ при $\delta=(h, 0)\to 0$. Используем теперь алгоритм о. п. к. Зададим число ω_δ ($\delta=(h, 0)$) так, чтобы выполнялось неравенство $1=\bar{\Omega}<\omega_\delta<\Omega_\delta\equiv\gamma(+0)=1+h^{-1/2}$. Тогда из графика функции $\gamma(\alpha)$, представленного на рис. 13, ясно, что алгоритм о. п. к. дает $\alpha_\delta=h^{3/2}$. Если теперь использовать правило отбора о. п. к. без учета числа λ_δ (см. § 2), то из рис. 14, на котором указаны функции $\beta(\alpha)$, $\pi(\alpha)$, следует, что выполнено неравенство (2.3): $I_\delta(z^{\alpha_2})>C\Pi_\delta(z^{\alpha_1})$.

Таким образом, выбор приближенного решения z^{α_δ} должен осуществляться по первой части правила отбора 2.2.2, т. е. нужно принять $z^{\alpha_\delta}=z^{\alpha_\delta}_{-}=h^{-1/2}$. Это также исключает сходимость $z^{\alpha_\delta}\to\bar{z}$ при $\delta\to 0$.

Имеется, однако, класс операторных уравнений вида (1.1), при приближенном решении которых с помощью алгоритмов обобщенных

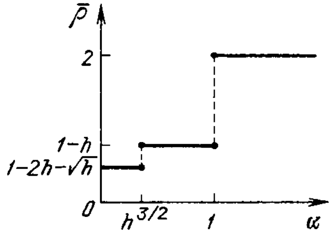

Рис. 12. График обобщенной невязки

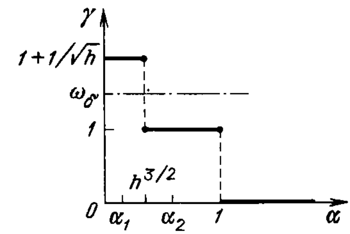

Рис. 13. График функции $\gamma(\alpha)$

принципов невязки, квазирешений и сглаживающего функционала можно не использовать величины λ_δ. Эти задачи характеризуются априорной информацией о разрешимости операторного уравнения на множестве D. В этом случае $\mu_0=0$. Определим тогда функции, аналогичные функциям (2.1), но не содержащие λ_δ:

$$\begin{gathered}\bar{\pi}(\alpha)\equiv Kf[\sigma+\psi(h,\Omega(z^\alpha))]\equiv K\bar{\Pi}_\delta(z^\alpha),\quad K=\mathrm{const}\geqslant 1,\\ \bar{\rho}(\alpha)\equiv\beta(\alpha)-\bar{\pi}(\alpha)\equiv P_\delta(z^\alpha)-I_\delta(z^\alpha)-K\bar{\Pi}_\delta(z^\alpha),\\ \bar{\varepsilon}(\alpha)\equiv\varphi(\alpha)-f\{[\psi(h,\Omega(z^\alpha))+\sigma]^p\}\equiv\bar{E}_\delta(z^\alpha),\\ 0<p<1.\end{gathered}\tag{1}$$

Сформулируем *модифицированный обобщенный принцип невязки* (м. о. п. н.) для нахождения Ω-оптимальных решений операторных уравнений [120].

1. В качестве параметра регуляризации α_δ выбирается решение уравнения с монотонно неубывающей функцией $\bar{\rho}(\alpha)$:

$$\bar{\rho}(\alpha)=0,\quad \alpha\geqslant 0.\tag{2}$$

2. а) В случае, если $\alpha_\delta>0$, приближенное решение z^{α_δ} выбирается из соответствующего множества Z^{α_δ} решений экстремальной задачи (1.9) по правилу отбора 2.3.1.

Если выполнено неравенство

$$I_\delta(z^{\alpha_2})\geqslant C\bar{\Pi}_\delta(z^{\alpha_1}),\tag{3}$$

то в качестве приближенного решения принимаем любой элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого справедливо неравенство

$$\bar{P}_\delta(z^{\alpha_\delta})\leqslant 0.\tag{4}$$

Например, можно выбрать $z^{\alpha_\delta}=z_-^{\alpha_\delta}$. Если же выполнено неравенство

$$I_\delta(z^{\alpha_2})\leqslant C\bar{\Pi}_\delta(z^{\alpha_1}), \tag{5}$$

то приближенным решением считается любой элемент $z^{\alpha_\delta}\in Z^{\alpha_\delta}$, для которого

$$\bar{P}_\delta(z^{\alpha_\delta})\geqslant 0. \tag{6}$$

В частности, можно принять $z^{\alpha_\delta}=z_+^{\alpha_\delta}$. В данном правиле отбора использованы величины C, $\alpha_{1,\,2}$, z^{α_1}, z^{α_2}, аналогичные фигурирующим в правиле 2.2.1.

б) Если $\alpha_\delta=0$, то в качестве приближенного решения можно взять любое решение экстремальной задачи: найти элемент $z_\delta^0\in Z_0$, для которого

$$\rho(A_h z_\delta^0,\ u_\sigma)=\inf\{\rho(A_h z,\ u_\sigma):\ z\in Z_0\}. \tag{7}$$

В дальнейшем будем предполагать, что выполнены предположения 1—3 из § 1. Спецификой разрешимых операторных уравнений вида (1.1) является то, что в алгоритмах о. п. н., о. п. к. можно использовать вспомогательные функции, на которые накладываются более слабые ограничения, чем предположение 4 из § 1. Будем считать, что функция $f(x)$ подчиняется предположению 5: $f(x)$ непрерывна и монотонно возрастает при $x\geqslant 0$. Поскольку значение функции $f(x)$ в точке $x=\mu_0=0$ известно, то без ограничения общности можно полагать $f(0)=0$.

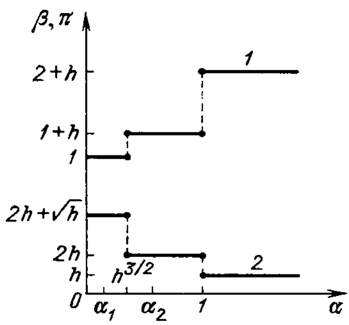

Рис. 14. Графики функций: 1 — $\beta(\alpha)$; 2 — $\pi(\alpha)$

Необходимо заметить, что решение уравнения (2) существует по крайней мере для достаточно малых $\|\delta\|$, если выполнено условие $Z_0\cap Z^*=\varnothing$. Действительно, при доказательстве теоремы 1.7.3 установлено, что требование $Z_0\cap Z^*=\varnothing$ влечет за собой неравенство $\nu_0\equiv\inf\{f[\rho(Az,\,u):\ z\in Z_0\}>f(\mu_0)=f(0)=0$ (см. теорему 1.7.4). Поэтому в силу сходимости $\nu_\delta\to\nu_0$ при $\|\delta\|\to 0$, обоснованной в теореме 1.7.3, и вследствие свойств функции ψ получим, что при малых $\|\delta\|$ выполнено неравенство $\bar{\rho}_\infty=\bar{\rho}(+\infty)=\nu_\delta-Kf\,[\sigma+\psi(h,\,\Omega^*)]>0$. Тогда по лемме 1.6.7 уравнение имеет решение α_δ, хотя это решение и может быть нулевым.

Касаясь задачи (7), отметим, что она разрешима вследствие выполнения предположений 1—3 из § 1 и по лемме 1.5.12.

Естественно, что замена предположения 4 из § 1 на предположение 5 приводит к необходимости специального обоснования сходимости алгоритма м. о. п. н. по иной схеме, нежели в § 1.7. Предположим, что параметр регуляризации $\alpha_\delta>0$ и приближенное решение z^{α_δ} получены при каждом δ ($\|\delta\|\leqslant\Delta_0$) по модифицированному обобщенному принципу невязки. В следующих леммах устанавливаются специальные свойства семейства $\{z^{\alpha_\delta}\}$.

Лемма 1. *Если элемент* z^{α_δ} *удовлетворяет условию* (6), *то для него выполнено неравенство* $\Omega(z^{\alpha_\delta})\leqslant\bar{\Omega}$. *Если же* $\Omega(z^{\alpha_\delta})>\bar{\Omega}$, *то для* z^{α_δ} *выполнено строгое неравенство* (4).

Доказательство проводится, как и в аналогичной лемме 1.6.7, с учетом исключения чисел λ_δ из алгоритма м. о. п. н.

Лемма 2. *Выполнено предельное соотношение*

$$\overline{\lim_{\delta\to 0}}\, \Omega(z^{\alpha_\delta})\leqslant\bar{\Omega}.$$

Доказательство. Если оказывается, что $\Omega(z^{\alpha_\delta})\leqslant\bar{\Omega}$, то лемма очевидна.

Пусть для каждого δ

$$\Omega(z^{\alpha_\delta})>\bar{\Omega}. \tag{8}$$

Тогда по лемме 1.5.4 с учетом неравенства (8) и равенства $f(0)=0$ получим из (1.5.7) при $\bar{\Omega}_\delta=\bar{\Omega}$

$$0<\Omega(z^{\alpha_\delta})-\bar{\Omega}\leqslant\{f[\sigma+\psi(h,\bar{\Omega})]-f(0)\}/\alpha_\delta=f[\sigma+\psi(h,\bar{\Omega})]/\alpha_\delta. \tag{9}$$

Возможны два варианта. В первом выполнено неравенство

$$\alpha_2=\alpha_\delta q\geqslant f^{1/2}[\sigma+\psi(h,\bar{\Omega})].$$

Из него и из (9) получим

$$0<\Omega(z^{\alpha_\delta})-\bar{\Omega}\leqslant qf^{1/2}[\sigma+\psi(h,\bar{\Omega})],$$

откуда, переходя к пределу при $\|\delta\|\to 0$ и учитывая непрерывность функций f, ψ, а также равенства $f(0)=0$, $\psi(0,\bar{\Omega})=0$, можно сделать вывод о том, что $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$.

Во втором варианте, когда

$$\alpha_2=\alpha_2(\delta)<f^{1/2}[\sigma+\psi(h,\bar{\Omega})],$$

ясно, что $\alpha_2(\delta)\to 0$ при $\|\delta\|\to 0$. Из неравенства (8) и леммы 1 можно понять, что элемент z^{α_δ} выбирается по первой части (3), (4) правила отбора 2.3.1. Тогда из (3), (8) выводится оценка

$$I_\delta(z^{\alpha_2})\geqslant Cf[\sigma+\psi(h,\Omega(z^{\alpha_1}))]\geqslant Cf[\sigma+\psi(h,\Omega(z^{\alpha_\delta}))]\geqslant Cf[\sigma+\psi(h,\bar{\Omega})].$$

Здесь использованы также свойства монотонности функций f,ψ, $\gamma(\alpha)=\Omega(z^\alpha)$.

По аналогии с переходом от (1.5.21) и (1.5.22) получим

$$\alpha_2\Omega(z^{\alpha_2})+Cf[\sigma+\psi(h,\Omega(z^{\alpha_1}))]\leqslant M^{\alpha_2}[z^{\alpha_2}]\leqslant$$
$$\leqslant M^{\alpha_2}[\bar{z}]\leqslant\alpha_2\bar{\Omega}+f[\sigma+\psi(h,\bar{\Omega})], \tag{10}$$

откуда следует неравенство

$$0\leqslant f[\sigma+\psi(h,\bar{\Omega})]/\alpha_\delta\leqslant q[\bar{\Omega}-\Omega(z^{\alpha_2})]/(C-1). \tag{11}$$

Следствием этой оценки оказываются два факта. Во-первых,

$$\Omega(z^{\alpha_2})\leqslant\bar{\Omega}. \tag{12}$$

Во-вторых, соотношения (11), (8), (9) дают

$$0<\Omega(z^{\alpha_\delta})-\bar{\Omega}\leqslant q\left[\bar{\Omega}-\Omega(z^{\alpha_2(\delta)})\right]/(C-1). \tag{13}$$

Кроме того, из (10) и из условия второго варианта на число α_2 получаем, принимая во внимание неотрицательность функции f,

$$0\leqslant I_\delta(z^{\alpha_2(\delta)})\leqslant\alpha_2(\delta)\bar{\Omega}+f[\sigma+\psi(h,\bar{\Omega})]\leqslant$$
$$\leqslant f^{1/2}[\sigma+\psi(h,\bar{\Omega})]\cdot\{\bar{\Omega}+f^{1/2}[\sigma+\psi(h,\bar{\Omega})]\}.$$

Это оценка после предельного перехода при $\|\delta\|\to 0$ дает сходимость

$$I_\delta(z^{\alpha_2(\delta)})\to 0=f(0)=f(\mu_0).$$

Полученная сходимость вместе с неравенством (12) позволяет применить к семейству $\{z^{\alpha_2(\delta)}\}$ замечание 1 к лемме 1.5.1, по которому оказывается, что $\Omega[z^{\alpha_2(\delta)}]\to\bar{\Omega}$ при $\|\delta\|\to 0$. Но тогда из (13) следует, что $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$. Лемма доказана.

Лемма 3. *Выполнено равенство*

$$\lim_{\delta\to 0} I_\delta(z^{\alpha_\delta})=\mu_0=0.$$

Доказательство проводится аналогично доказательству лемм 1.5.8, 1.5.9 с учетом равенства $\mu_0=0$.

Теорема 1. *Пусть выполнены предположения* 1—3 *из* § 1 *и предположение* 5, *а последовательность* $\{\delta_n\}$ *сходится к нулю при* $n\to\infty$. *Если* $\alpha_n\equiv\alpha_{\delta_n}>0$, $z_n\equiv z^{\alpha_n}$ *определены по м.о.п.н., то* $z_n\xrightarrow{\tau}\bar{Z}$ *и* $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Доказательство основано на результатах лемм 2, 3, которые являются условиями следствия 1.5.1. Из этого следствия и получаются искомые утверждения.

Теорема 1 показывает, что м.о.п.н. представляет собой регуляризирующий алгоритм для решения операторных уравнений вида (1.1) при условии $\alpha_\delta>0$.

Отметим частный случай алгоритма м.о.п.н., получающийся при $K=C$ и используемый при точном задании оператора задачи (1.1), т. е. при $h=0$. В этом случае уравнение (2) для выбора параметра регуляризации α_δ $(\delta=(0,\sigma))$ принимает вид

$$f[\rho(Az^\alpha,u_\sigma)]=Cf(\sigma).$$

Поэтому вследствие монотонного неубывания функции $\beta(\alpha)$ для α_2 $(\alpha_2=\alpha_\delta q>\alpha_\delta)$ выполнено неравенство

$$\beta(\alpha_2)=f[\rho(Az^{\alpha_2},u_\sigma)]\geqslant\beta(\alpha_\delta+0)\geqslant Cf(\sigma)=C\Pi_\delta(z^{\alpha_1}),$$

которое совпадает с условием (3). Таким образом, в рассматриваемом случае правило отбора алгоритма м.о.п.н. значительно упрощается, так как приближенное решение z^{α_δ} всегда выбирается так, чтобы выполнялось неравенство (4):

$$f[\rho(Az^{\alpha_\delta},u_\sigma)]\leqslant Cf(\sigma).$$

Исследуем теперь случай, когда алгоритм м.о.п.н. формально дает $\alpha_\delta = 0$.

Теорема 2. *Пусть выполнены предположения* 1—3 *из* § 1 *и предположение* 5, *а функция* $\psi(h, \Omega)$ *возрастает по второму аргументу при каждом* $h > 0$. *Если для последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n \to \infty$ *и такой, что* $h_n \neq 0$, *при использовании м.о.п.н. оказывается выполненным равенство* $\alpha_n \equiv \alpha_{\delta_n} = 0$ *и по правилу отбора* 2.3.1, б) *приближенное решение* $z_n = z^0_{\delta_n}$ *выбрано как решение задачи* (7), *то* $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n \to \infty$.

Доказательство. Заметим прежде всего, что условие теоремы $\alpha_n = 0$, выполненное для любого номера n, ведет к непустоте пересечения множеств Z^* и Z_0. Установить это можно, используя в основном технику доказательства теоремы 1.7.3. При этом обосновывается с учетом неравенства $\mu_\delta \geqslant 0$, что из условия $Z^* \cap Z_0 = \varnothing$ при достаточно малых $\|\delta\|$ следует справедливость неравенств

$$\bar{\rho}^\delta_0 \equiv \bar{\rho}_\delta(+0) = f(\mu_\delta) - Kf[\sigma + \psi(h, \Omega_\delta)] < 0,$$
$$\bar{\rho}^\delta_\infty \equiv \bar{\rho}_\delta(+\infty) = \nu_\delta - Kf[\sigma + \psi(h, \Omega^*)] > 0.$$

Однако тогда по лемме 1.6.7 оказывается, что $\alpha_n > 0$ по крайней мере при $n \geqslant n_0 = \mathrm{const}$. Это противоречие доказывает, что $Z^* \cap Z_0 \neq \varnothing$, и поэтому

$$\inf\{\rho(Az, \bar{u}): z \in D\} = \inf\{\rho(Az, \bar{u}): z \in Z_0\}.$$

В свою очередь последнее равенство означает, что $\bar{Z} \subset Z_0$. Тогда из результатов теоремы 1.10.3 следует, что последовательность $\{z_n\}$ решений задачи (7) при $\delta = \delta_n$ обладает свойствами сходимости: $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega} = \Omega^*$ при $n \to \infty$.

Итак, и в случае, когда $\alpha_\delta = 0$, алгоритм м.о.п.н. дает устойчивый метод решения совместного операторного уравнения (1.1).

Сформулируем теперь *модифицированный обобщенный принцип квазирешений* (м.о.п.к.). Пусть задана величина ω_δ ($\omega_\delta > \Omega^*$) — оценка для $\bar{\Omega}$.

1. В качестве параметра регуляризации $\alpha_\delta \geqslant 0$ выбирается решение уравнения с монотонно не возрастающей функцией $\gamma(\alpha)$:

$$\gamma(\alpha) = \omega_\delta, \quad \alpha \geqslant 0. \tag{14}$$

2. а) Если оказывается, что $\alpha_\delta > 0$, приближенное решение z^{α_δ} выбирается по правилу отбора 2.3.2.

Согласно этому правилу, при выполнении неравенства (3) в качестве приближенного решения принимается элемент $z^{\alpha_\delta} \in Z^{\alpha_\delta}$, для которого справедливо требование

$$\Omega(z^{\alpha_\delta}) \geqslant \omega_\delta. \tag{15}$$

Так, например, можно взять $z^{\alpha_\delta} = z^{\alpha_\delta}_-$. Если же выполнено соотношение (5), то приближенное решение z^{α_δ} выбираем из Z^{α_δ} так, чтобы для него было верным условие

$$\Omega(z^{\alpha_\delta}) \leqslant \omega_\delta. \tag{16}$$

В частности, можно положить, что $z^{\alpha_\delta} = z^{\alpha_\delta}_+$.

б) Если параметр регуляризации α_δ равен нулю, то в качестве приближенного решения можно принять любой элемент z_δ^0, для которого

$$\rho(A_h z_\delta^0,\ u_\sigma) = \inf\{\rho(A_h z,\ u_\sigma):\ z \in D\}. \tag{17}$$

Из теоремы 1.8.1 и леммы 1.6.7 ясно, что при условии $\omega_\delta > \Omega^*$ решение уравнения (14) существует, причем при $\omega_\delta < \Omega_\delta \equiv \gamma(+0)$ оно положительно. Из предположений 1—3 из § 1 и из теоремы 1.10.2 ясно также, что задача (17) разрешима в случае, когда $\alpha_\delta = 0$.

По аналогии с доказательством леммы 2 и с привлечением леммы 1.5.11, в которой следует положить $\bar{\Omega}_\delta \equiv \omega_\delta$, можно установить справедливость следующей леммы.

Лемма 4. *Если величина ω_δ удовлетворяет условиям $\bar{\Omega} \leqslant \omega_\delta < \Omega_\delta$, $\omega_\delta \to \bar{\Omega}$ при $\|\delta\| \to 0$, то справедливы соотношения*

$$\varlimsup_{\delta \to 0} \Omega(z^{\alpha_\delta}) \leqslant \bar{\Omega}, \quad \lim_{\delta \to 0} I_\delta(z^{\alpha_\delta}) = 0.$$

Результат этой леммы позволяет утверждать, что для приближений z^{α_δ}, полученных по алгоритму м.о.п.н. в случае, когда $\alpha_\delta > 0$, будет справедлива теорема сходимости 1. В свою очередь для случая $\alpha_\delta = 0$ из теорем 1.10.2, 1.10.3 можно получить теорему сходимости приближений z_δ^0.

Теорема 3. *Предположим, что выполнены предположения* 1—3 *из* § 1 *и предположение* 5. *Если для сходящейся к нулю при $n \to \infty$ последовательности $\{\delta_n\}$ $(h_n \neq 0)$ оказывается, что при использовании алгоритма м.о.п.к. выполнено равенство $\alpha_n \equiv \alpha_{\delta_n} = 0$ и приближенное решение $z_n \equiv z_{\delta_n}^0$ выбирается как решение задачи* (17), *то $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ при $n \to \infty$.*

Таким образом, модифицированный обобщенный принцип квазирешений представляет собой регуляризующий алгоритм для разрешимых операторных уравнений (1.1).

Сформулируем *модифицированный обобщенный принцип сглаживающего функционала* (м.о.п.с.ф.). Будем при этом считать, что $\Omega^* \geqslant 0$.

1. Параметр регуляризации $\alpha_\delta \geqslant 0$ выбирается как решение уравнения с монотонной функцией

$$\bar{\varepsilon}(\alpha) = 0. \tag{18}$$

2. а) Если $\alpha_\delta > 0$, то приближенное решение z^{α_δ} выбирается из множества Z^{α_δ} решений задачи (1.9) при $\alpha = \alpha_\delta$ по правилу отбора 2.3.3, т. е. как любой элемент $z^{\alpha_\delta} \in Z^{\alpha_\delta}$, для которого выполнено неравенство $\bar{E}_\delta(\dot{z}^{\alpha_\delta}) \geqslant 0$. Так, можно принять $z^{\alpha_\delta} = z_+^{\alpha_\delta}$.

б) Если $\alpha_\delta = 0$, то в качестве приближенного решения можно взять любое решение задачи (7).

Как и для алгоритма м.о.п.н., решение уравнения (18) существует по крайней мере при достаточно малых $\|\delta\|$, если $Z_0 \cap Z^* = \varnothing$. Путем почти дословного повторения доказательства теоремы 1.9.5 с заменой в соответствующих формулах величин J^*, J_δ^*, λ_δ на нули устанавливается.

Теорема 4. *Пусть выполнены предположения* 1—4 *из* §1, *а также требования* $Z^*\bigcap Z_0=\varnothing$, $\Omega^*\geqslant 0$: *Если* $\alpha_\delta>0$ *определено по м.о.п.с.ф., то* $\alpha_\delta\to 0$ *при* $\|\delta\|\to 0$.

Из теоремы 1.9.6 аналогично следует

Теорема 5. *Предположим, что кроме условий теоремы* 4 *выполнены также дополнительные условия на функцию* $\psi(h,\Omega)$: *при каждом* $h>0$ *она монотонно возрастает по второму аргументу Если для последовательности* $\{\delta_n\}$, *сходящейся при* $n\to\infty$ *к нулю величины* $\alpha_n\equiv\alpha_{\delta_n}>0$, $z_n\equiv z^{\alpha_n}$ *определены по м.о.п.с.ф., то* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

В особом случае $\alpha_\delta=0$ из теоремы 1.10.3 вытекает

Теорема 6. *Если при выполнении предположений* 1—4 *из* §1 *и условия теоремы* 5 *на функцию* ψ *для последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n\to\infty$ *и такой, что* $h_n\neq 0$, *при использовании м.о.п.с.ф. оказывается выполненным равенство* $\alpha_n\equiv\alpha_{\delta_n}=0$, *то приближенные решения* $z_n\equiv z^0_{\delta_n}$*—решения задачи* (7′)*—обладают свойствами* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Подводя итоги сказанному, следует отметить, что априорная информация о разрешимости задачи (1.1) позволяет упростить алгоритмы о.п.н., о.п.к., о.п.с.ф. путем исключения из них величины λ_δ, которая требует специальной процедуры вычисления. Кроме того, оказывается, что модифицированные о.п.н., о.п.к. могут быть обоснованы для более широкого класса вспомогательных функций f, нежели обычные версии этих алгоритмов, рассмотренные в § 2.

Для случая линейных операторных уравнений м.о.п.н. рассматривался в [88, 104, 186, 187].

§ 4. Некоторые частные случаи применимости алгоритмов о.п.н., о.п.к., о.п.с.ф. для решений операторных уравнений

В § 1.11 приведены для общего случая решения некорректных экстремальных задач примеры часто встречающихся в приложениях функционалов Ω, мер аппроксимации Ψ, конкретных вариантов условий, обеспечивающих применимость алгоритмов о.п.н., о.п.к., о.п.с.ф. Аналогичные условия возникают и для случая решения вариационным методом операторных уравнений, причем в этом случае общие требования из § 1 могут быть значительно детализированы.

Рассмотрим предположение 1 из § 1: для любого элемента u_0 из метрического пространства U метрика $\rho(u, u_0)$ t-секвенциально полунепрерывна снизу в U по аргументу u. На практике весьма часто встречается случай, когда U—нормированное пространство. Тогда можно положить $\rho(u, u_0)=\|u-u_0\|$ и для выполнения предположения 1 в качестве t выбрать топологию, порождаемую нормой в U. Если же U—банахово пространство, то в качестве t можно взять также топологию, порождающую слабую сходимость в U. В силу слабой полунепрерывности снизу нормы в банаховом пространстве (см. § 0.3) предположение 1 также будет выполнено. Сказанное справедливо в случае U — гильбертово пространство. Возможны случаи, в которых топология t не связана непосредственно с сильной или слабой сходимостью в U. Приведем соответствующие примеры, когда предположение 1 также выполнено.

Пример 1. Предположим, что $U=H^{\mu}[a, b]$ — пространство функций, удовлетворяющих на отрезке $[a, b]$ условию Гёльдера с показателем μ $(0<\mu\leqslant 1)$. Если тогда t — топология равномерной сходимости в U, то, как показано в § 1.11, норма $\|u\|_{\mu}$ в пространстве $H^{\mu}[a, b]$ является t-секвенциальной полунепрерывной снизу на U.

Пример 2. Пусть $U=V[a, b]$ — пространство функций с ограниченной вариацией на $[a, b]$ и с нормой

$$\|u\|_V=|u(a)|+\bigvee_a^b(u).$$

В качестве топологии t возьмем топологию, порождающую поточечную сходимость ограниченных в U последовательностей: последовательность $\{u_n(s)\}\subset V[a, b]$ t-сходится к $u_0(s)\in V[a, b]$, если $u_n(s)\to u_0(s)$ при $n\to\infty$ для любого $s\in[a, b]$ и $\|u_n\|_V\leqslant C_0=\text{const}$. Тогда норма $\|\cdot\|_V$ будет t-полунепрерывна снизу в U (см. лемму 4.1.1).

Заметим, что в приведенных примерах пространство U нерефлексивное [74].

Рассмотрим теперь в комплексе несколько конкретных вариантов предположений 1, 2 из § 1.

Условие А. U — метрическое пространство с топологией t, порожденной метрикой ρ; Z — банахово пространство с топологией τ слабой сходимости; множество D слабо замкнуто в Z; операторы A, A_h усиленно непрерывны из Z в U (см. § 0.3); функционал $\Omega(z)$ слабо полунепрерывен снизу (слабо непрерывен) на множестве D; множества $\Omega_C=\{z\in D\colon\ \Omega(z)\leqslant C\}$ слабо компактны, если не являются пустыми.

В частности, условия усиленной непрерывности операторов A, A_h выполнены, если они представляют собой линейные вполне непрерывные операторы из Z в U. Важным примером нелинейных усиленно непрерывных операторов A, A_h являются операторы вида $A=A^0B$, $A_h=A_h^0B$, где операторы A^0, A_h^0 непрерывны из банахова пространства V в U, а B — линейный вполне непрерывный оператор, действующий из Z в V. В частности, оператор B может быть оператором компактного вложения пространства Z в V.

Приведем конкретные примеры усиленно непрерывных операторов.

Пример 3. Зададим функции

$$K_0(x, s)\in C\{[c, d]\times[a, b]\},\ f_0(s, z)\in C\{[a, b]\times\mathbf{R}\}$$

и введем оператор

$$A[x, z(s)]=\int_a^b K_0(x, s)f_0[s, z(s)]ds,\ x\in[c, d].$$

Действующий из пространства $C[a, b]$ в пространство $C[c, d]$ оператор A будет усиленно непрерывным. Действительно, пусть имеется последовательность функций $\{z_n(s)\}\subset C[a, b]$, которая слабо сходится в C к $z_0(s)\in C[a, b]$. Это означает (см. теорему 0.4.4), что $|z_n(s)|\leqslant M=\text{const}$ для любого номера n и $z_n(s)\to z_0(s)$ для любой точки $s\in[a, b]$. Тогда

$$f_0[s, z_n(s)]\to f_0[s, z_0(s)] \quad \forall s\in[a, b],$$
$$|f_0(s, z_n(s))|\leqslant F_0\equiv\max\{|f_0(s, z)|: s\in[a, b], |z|\leqslant M\}=\text{const}.$$

Отсюда по теореме Лебега (см. теорему 0.4.8) получим

$$|A[x, z_n(s)]-A[x, z_0(s)]|\leqslant$$
$$\leqslant\int_a^b|K_0(x, s)|\cdot|f_0[s, z_n(s)]-f_0[s, z_0(s)]|\,ds\leqq$$
$$\leqslant\|K_0\|_C\int_a^b|f_0[s, z_n(s)]-f_0[s, z_0(s)]|\,ds\to 0, \quad n\to\infty.$$

Но тогда последовательность $\{A[x, z_n(s)]\}$ равномерно сходится к $A[x, z_0(s)]$ на отрезке $[c, d]$ при $n\to\infty$. Таким образом, оператор A преобразует всякую слабо сходящуюся в $C[a, b]$ последовательность в сильно сходящуюся последовательность функций из $C[c, d]$.

Пример 4. Будем считать, что задана функция $K(x, s, z)\in\in C\{[c, d]\times[a, b]\times\mathbf{R}\}$, обладающая следующим свойством: для каждых $x\in[c, d]$, $s\in[a, b]$ и любой пары $z_1, z_2\in\mathbf{R}$

$$|K(x, s, z_1)-K(x, s, z_2)|\leqslant|K_0(x, s)|\cdot|f_0(s, z_1)-f_0(s, z_2)|,$$

где K_0, f_0 удовлетворяют условиям примера 3. Тогда оператор

$$A[x, z(s)]=\int_a^b K(x, s, z(s))\,ds, \quad x\in[c, d],$$

рассматриваемый из $C[a, b]$ в $C[c, d]$, также усиленно непрерывен. Доказательство этого факта проводится по той же схеме, что и в примере 3.

Другие примеры и достаточные условия усиленной непрерывности нелинейных операторов, использующие их потенциальность, дифференцируемость и полную непрерывность производной, приведены в [26], а также в [32, 48].

Условие В. U, Z — банаховы пространства с топологиями слабой сходимости; множество D слабо замкнуто в Z; операторы A, A_h слабо непрерывны из Z в U; функционал $\Omega(z)$ удовлетворяет тем же требованиям, что и в условиях А.

Условие слабой непрерывности операторов A, A_h выполняется, например, в случае, если эти операторы являются линейными и ограниченными при действии из банахова пространства Z в банахово пространство U. Другой пример слабо непрерывных операторов получается, если рассматривать операторы вида $A=A^0B$, $A_h=A_h^0B$, где B — линейный, вполне непрерывный оператор, действующий из Z в банахово пространство V, а A^0, A_h^0 — деминепрерывные операторы, действующие из V в U. Содержательные примеры подобных операторов приведены в [26, 48]. Кроме того, следует заметить, что усиленно непрерывные операторы A, A_h также будут слабо непрерывными.

Условия C. U, Z — банаховы пространства; t — топология слабой сходимости в U; τ — топология сильной сходимости в Z; множество D замкнуто в Z; операторы A, A_h деминепрерывны из Z в U; $\Omega(z)$ — полунепрерывный снизу (непрерывный) на D функционал; множество Ω_C компактно в Z для всякого C, для которого оно непусто.

Условия D. U, Z — метрические пространства с топологиями, порождаемыми метриками; D замкнуто в Z; A, A_h — непрерывные из Z в U операторы; требования на функционал Ω те же, что и в условиях C.

Тот факт, что условия A — D представляют собой частные случаи предположений 1, 2 из § 1, вытекает из материала, изложенного в § 0.3.

В соответствии с теоремами 2.3, 2.6, 2.9, а также с теоремами 3.1—3.3, 3.5, 3.6 условия A, B гарантируют, что приближенные решения z^{α_δ}, которые получены с помощью алгоритмов о.п.н., о.п.к., о.п.с.ф. или же с помощью м.о.п.н., м.о.п.к., м.о.п.с.ф., будут слабо сходиться к множеству $\bar{Z}$ Ω-оптимальных квазирешений уравнения (1.1), причем при этом $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$ для $\|\delta\|\to 0$. Весьма важен для приложений частный случай, когда Z — рефлексивное банахово пространство с H-свойством, а функционал $\Omega(z)$ задается в виде $\Omega(z)=g(\|z\|)$, где $g(x)$ — непрерывная и возрастающая при $x\geqslant 0$ функция, обладающая также свойством $g(+\infty)=+\infty$. Тогда сходимость $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$ при $\|\delta\|\to 0$ влечет за собой сходимость $\|z^{\alpha_\delta}\|\to\|\bar{z}\|=g^{-1}(\Omega)$ $(\bar{z}\in\bar{Z})$. Поскольку z^{α_δ} при этом слабо сходится к $\bar{Z}$, то в силу H-свойства элементы z^{α_δ} обладают также и свойством сильной сходимости к $\bar{Z}$ при $\|\delta\|\to 0$.

Что касается условий C, D, то они по указанным теоремам обеспечивают сильную сходимость приближений к множеству $\bar{Z}$, а также по-прежнему сходимость $\Omega(z^{\alpha_\delta})\to\bar{\Omega}$ при $\|\delta\|\to 0$.

Рассмотрим более подробно важный частный случай задачи (1.1), когда операторы A, A_h линейные.

Лемма 1. 1. *Пусть A — линейный ограниченный оператор, действующий из нормированного пространства Z в нормированное пространство U. Предположим также, что D — выпуклое множество из Z, а функция $f(x)$, определенная при $x\geqslant 0$, выпуклая и неубывающая. Тогда для любого фиксированного $u\in U$ функционал $I(z)=f(\|Az-u\|)$ является выпуклым на D.*

2. *Если $g(x)$ — строго выпуклая, возрастающая и непрерывная при $x\geqslant 0$ функция, а Z — строго нормированное пространство, то функционал $g(\|z\|)$ будет строго выпуклым на любом выпуклом множестве $D\subset Z$.*

Доказательство. 1. Достаточно убедиться, что для любой пары z_1, $z_2\in D$ выполнено неравенство $I[(z_1+z_2)/2]\leqslant\{I(z_1)+I(z_2)\}/2$. Чтобы убедиться в этом, используем вытекающие из линейности оператора A и свойств нормы неравенство

$$\left\|A\left(\frac{z_1+z_2}{2}\right)-u\right\|=\left\|\frac{(Az_1-u)+(Az_2-u)}{2}\right\|\leqslant\frac{\|Az_1-u\|+\|Az_2-u\|}{2}$$

и воспользуемся неубыванием и выпуклостью функции f в следующей выкладке:

$$I\left(\frac{z_1+z_2}{2}\right)=f\left[\left\| A\left(\frac{z_1+z_2}{2}\right)-u\right\|\right]\leqslant f\left(\frac{\| Az_1-u\|+\| Az_2-u\|}{2}\right)\leqslant$$

$$\leqslant\frac{f(\| Az_1-u\|)+f(\| Az_2-u\|)}{2}=\frac{I(z_1)+I(z_2)}{2}\quad \forall z_1,\ z_2\in D$$

2. Рассмотрим сначала случай произвольных элементов z_1, $z_2 \in D$ с различной нормой. Тогда из возрастания и строгой выпуклости функции g с учетом выпуклости нормы получим

$$g(\|\alpha z_1+(1-\alpha)z_2\|)\leqslant g(\alpha\| z_1\|+(1-\alpha)\| z_2\|)\leqslant \alpha g(\| z_1\|)+(1-\alpha)g(\| z_2\|)$$

для любых чисел α $(0\leqslant\alpha\leqslant 1)$, причем равенство в последнем неравенстве возможно вследствие строгой выпуклости функции g лишь при $\alpha=0$ или $\alpha=1$.

Пусть теперь $\| z_1\|=\| z_2\|$. Если при этом элементы z_1, z_2 линейно независимы, то в силу строгой нормированности пространства Z выполнено неравенство

$$\|\alpha z_1+(1-\alpha)z_2\|<\alpha\| z_1\|+(1-\alpha)\| z_2\|$$

для $0<\alpha<1$. Поэтому из-за возрастания функции g для $0<\alpha<1$ получается, что

$$g(\|\alpha z_1+(1-\alpha)z_2\|)<g(\alpha\| z_1\|+(1-\alpha)\| z_2\|)\leqslant\alpha g(\| z_1\|)+(1-\alpha)g(\| z_2\|).$$

Если же z_1, $z_2\in D$ линейно зависимы, причем $\| z_1\|=\| z_2\|$, $z_1\neq z_2$, то тогда $z_1=-z_2$ и

$$g(\|\alpha z_1+(1-\alpha)z_2\|)=g(\|(2\alpha-1)z_1\|)=g(|2\alpha-1|\,\| z_1\|)\leqslant$$

$$\leqslant\alpha g(\| z_1\|)+(1-\alpha)g(\| z_2\|)=g(\| z_1\|).$$

Вследствие возрастания функции g равенство здесь возможно, если лишь $|2\alpha-1|=1$, т. е. при $\alpha=0$ или $\alpha=1$.

Резюмируя результаты всех рассмотренных случаев, можно сделать вывод о том, что неравенство

$$g(\|\alpha z_1+(1-\alpha)z_2\|)\leqslant\alpha g(\| z_1\|)+(1-\alpha)g(\| z_2\|),$$

выполненное при всех α $(0\leqslant\alpha\leqslant 1)$, $z_1\neq z_2$, может превращаться в равенство лишь при $\alpha=0$ или $\alpha=1$. Это и означает строгую выпуклость функционала $g(\| z\|)$.

Пусть теперь выполнены

Условия Е. Z — рефлексивное банахово пространство с топологией слабой сходимости τ; U — нормированное пространство; D — выпуклое замкнутое множество в Z; A, A_h — линейные ограниченные операторы из Z в U; функционал $\Omega(z)$ имеет вид $\Omega(z)=g(\| z\|)$, где функция $g(x)$ монотонно возрастает и непрерывна при $x\geqslant 0$, причем $g(+\infty)=+\infty$; функция $f(x)$ выпукла при $x\geqslant 0$ и $f(x)\in\mathcal{F}^m$ $(m\geqslant 1)$.

Для линейных операторов выполнено также условие аппроксимации

$$\| Az-A_hz\|_U\leqslant\| A-A_h\|\cdot\| z\|_Z\leqslant h\| z\|=hg^{-1}[\Omega(z)]\equiv\psi(h,\ \Omega(z))\quad \forall z\in Z.$$

Здесь h — известная оценка величины $\| A-A_h \|$, характеризующей погрешность задания оператора. Сформулированные условия гарантируют возможность применения алгоритмов о.п.н., о.п.к., о.п.с.ф. для нахождения нормальных квазирешений линейных уравнений типа (1.1). Действительно, по лемме 1 функционал $I_\delta(z)=f(\| A_h z-u_\delta \|)$ будет выпуклым на множестве D, а в силу непрерывности оператора A_h еще и непрерывным на D. Тогда, как отмечено в теореме 0.3.8, этот функционал будет также и слабо полунепрерывным снизу на множестве D. Функционал $\Omega(z)$ будет также слабо полунепрерывным снизу на D. Действительно, норма в рефлексивном банаховом пространстве Z обладает свойством слабой полунепрерывности снизу: для любой последовательности $\{z_n\}$, слабо сходящейся при $n\to\infty$ к элементу z_0, справедливо неравенство

$$\varliminf_{n\to\infty} \| z_n \| \geqslant \| z_0 \|. \tag{1}$$

Пусть теперь $\{z_n\}\subset D$ — произвольная, слабо сходящаяся к элементу z_0 последовательность. Тогда по условиям E множество D, являясь выпуклым и замкнутым, будет слабо замкнуто (см. теорему 0.3.7). Поэтому $z_0\in D$, и из неравенства (1) с учетом возрастания и непрерывности функции g получим

$$\varliminf_{n\to\infty} g(\| z_n \|)\geqslant g(\| z_0 \|).$$

Это и доказывает слабую полунепрерывность снизу на D функционала Ω.

Множество

$$\Omega_C=\{z\in D:\ g(\| z \|)\leqslant C\}=\{z\in D:\ \| z \|\leqslant g^{-1}(C)\}$$

слабо замкнуто и слабо компактно по свойствам шара в рефлексивном банаховом пространстве Z и с учетом свойств множества D и функции g.

Тем самым основные условия, обеспечивающие применимость алгоритмов обобщенных принципов для решения линейных задач (1.1), вытекают из условий E. При этом алгоритмы дают слабую сходимость приближений z^{α_δ} к множеству $\bar{Z}$ нормальных квазирешений задачи (1.1) и, кроме того, $\Omega(z^{\alpha_\delta})=g(\| z^{\alpha_\delta} \|)\to\bar{\Omega}=g(\| \bar{z} \|)$ при $\| \delta \|\to 0$. Последняя сходимость влечет в силу непрерывности и возрастания функции g сходимость норм $\| z^{\alpha_\delta} \|\to\| \bar{z} \|=g^{-1}(\bar{\Omega})(\bar{z}\in\bar{Z})$. Если в дополнение к условиям E пространство Z обладает H-свойством, то из слабой сходимости и сходимости норм следует сильная сходимость приближений к $\bar{Z}$ при $\| \delta \|\to 0$.

Отметим, что условия E не гарантируют единственности экстремали сглаживающего функционала при каждом $\alpha>0$. Можно, однако, сформулировать условия, обеспечивающие это требование.

Условия F. Z — локально равномерно выпуклое рефлексивное банахово пространство; τ, U, D, f, A, A_h — те же, что и в условиях E; $\Omega(z)=g(\| z \|)$, где функция $g(x)$ непрерывна, монотонно возрастает и строго выпукла при $x\geqslant 0$.

Как указано в § 0.3 (теорема 0.3.2), локально равномерно выпуклое рефлексивное банахово пространство является строго нормированным и обладает H-свойством. Поэтому в силу леммы 1 сглаживающий функционал

$$M^\alpha[z]=\alpha g(\|z\|)+f(\|A_h z-u_\sigma\|), \quad z\in D, \tag{2}$$

будет строго выпуклым, и поэтому по теоремам 0.2.1, 0.3.5—0.3.8 задача его минимизации на выпуклом замкнутом множестве D имеет единственное решение. Это гарантирует непрерывность вспомогательных функций (2.1), и, следовательно, уравнения (2.2) (или (3.2)), (2.5), (2.6) (или (3.18)), используемые при выборе параметра регуляризации и принимающие в данном случае вид

$$\|A_h z^\alpha-u_\sigma\|=\psi(h, g(\|z^\alpha\|))+\sigma+\lambda_\delta, \tag{2.2'}$$

$$\|A_h z^\alpha-u_\sigma\|=\psi(h, g(\|z^\alpha\|))+\sigma, \tag{3.2'}$$

$$g(\|z^\alpha\|)=\omega_\delta, \tag{2.5'}$$

$$\alpha g(\|z^\alpha\|)+f(\|A_h z^\alpha-u_\sigma\|)=f\{\lambda_\delta+[\psi(h, g(\|z^\alpha\|))+\sigma]^p\}, \tag{2.6'}$$

$$\alpha g(\|z^\alpha\|)+f(\|A_h z^\alpha-u_\sigma\|)=f\{[\sigma+\psi(h, g(\|z^\alpha\|))]^p\}, \tag{3.18'}$$

имеют или положительные обычные решения α_δ, или нулевое обобщенное решение. В частности, если $\alpha_\delta>0$, то отпадает необходимость использования правил отбора, и в качестве приближенного решения можно взять элемент z^{α_δ} — единственную экстремаль сглаживающего функционала (2), отвечающую значению $\alpha=\alpha_\delta$. Согласно условиям F, эти приближения будут сильно сходиться к $\bar{Z}$ при $\|\delta\|\to 0$.

В заключение обратимся более подробно к случаю использования обобщенного принципа невязки для решения линейного операторного уравнения (1.1) в гильбертовых пространствах Z, U $(D=Z)$. При этом будем считать, что $f(x)=g(x)=x^2$. Тогда

$$M^\alpha[z]=\alpha\|z\|^2+\|A_h z-u_\sigma\|^2, \quad \alpha>0, \tag{3}$$

и, как только что было отмечено, задача минимизации этого функционала в Z имеет единственное решение z^α при любом $\delta=(h, \sigma)$. Для нахождения элемента z^α в силу сильной выпуклости функционала (3) справедливо необходимое и достаточное условие минимума [202]:

$$\alpha z^\alpha+A_h^* A_h z^\alpha=A_h^* u_\sigma. \tag{4}$$

Здесь A_h^* — оператор, сопряженный с A_h.

Учитывая, что мера аппроксимации ψ для изучаемого случая имеет вид $\psi(h, \Omega(z))=h\|z\|$, можно определить вспомогательные функции (2.1)—(3.1) следующим образом (рис. 15):

$$\varphi(\alpha)=\alpha\|z^\alpha\|^2+\|A_h z^\alpha-u_\sigma\|^2, \quad \gamma(\alpha)=\|z^\alpha\|^2,$$

$$\beta(\alpha)=\|A_h z^\alpha-u_\sigma\|^2; \tag{5}$$

$$\rho(\alpha)=\beta(\alpha)-(\sigma+\lambda_\delta+h\|z^\alpha\|)^2, \quad \bar{\rho}(\alpha)=\beta(\alpha)-(\sigma+h\|z^\alpha\|)^2. \tag{5'}$$

Тогда условия выбора (2.2′), (3.2′) параметра регуляризации по обобщенному принципу невязки конкретизируются следующим образом:

$$\rho(\alpha) \equiv \| A_h z^\alpha - u_\sigma \|^2 - (\sigma + \lambda_\delta + h \| z^\alpha \|)^2 = 0, \qquad (2.2'')$$

$$\bar{\rho}(\alpha) \equiv \| A_h z^\alpha - u_\sigma \|^2 - (\sigma + h \| z^\alpha \|)^2 = 0. \qquad (3.2'')$$

Входящие в эти выражения функции, так же как и другие вспомо-

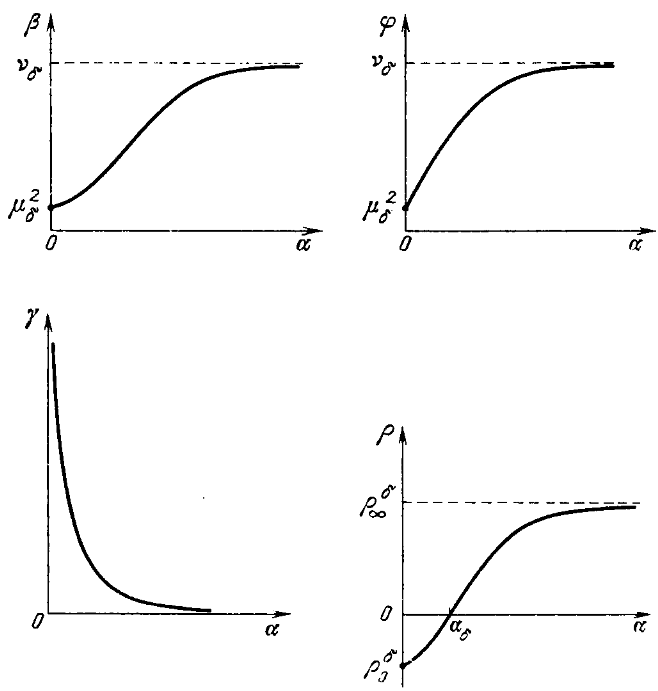

Рис. 15. Типичные графики вспомогательных функций в линейном случае

гательные функции (5), непрерывны при $\alpha > 0$ (см. следствие 1.6.1). Из результатов леммы 1.11.2 и теоремы 1.11.1 следует, что уравнения (2.2″), (3.2″) имеют единственные обычные решения $\alpha_\delta > 0$, $\bar{\alpha}_\delta > 0$ при условиях $\| u_\sigma \| > \sigma + \lambda_\delta$, $\| u_\sigma \| > \sigma$. Для численного нахождения этих решений существенную роль играет

Лемма 2. 1. *Функции* (5) *непрерывно дифференцируемы при* $\alpha > 0$:

$$\varphi'(\alpha) = \gamma(\alpha), \quad \gamma'(\alpha) = -\langle (A_h^* A_h + \alpha E)^{-1} z^\alpha, \ z^\alpha \rangle,$$

$$\beta'(\alpha) = -\alpha \gamma'(\alpha).$$

Если соответственно выполнены условия

$$\| u_\sigma \| > \sigma + \lambda_\delta, \quad \| u_\sigma \| > \sigma, \qquad (6)$$

то функции (5′) *также непрерывно дифференцируемы*:

$$\rho'(\alpha) = -\gamma'(\alpha)\left[\alpha + h^2 + h(\sigma + \lambda_\delta)/\sqrt{\gamma(\alpha)}\right],$$
$$\bar{\rho}'(\alpha) = -\gamma'(\alpha)\left[\alpha + h^2 + h\sigma/\sqrt{\gamma(\alpha)}\right].$$

2. *При выполнении условий* (6) *функции* (5), (5′) *строго монотонны*
3. *Функция* $\beta(1/\lambda)$ *выпукла по* λ *при* $\lambda > 0$.

Здесь $\langle\, ,\rangle$ — *скалярное произведение в* Z.

Дифференцируемость функций (5), (5′) и выпуклость функции $\beta(1/\lambda)$ отмечены в работах [160, 164, 206, 226].

Как уже было сказано в § 1.14, для реализации о.п.н. важную роль играет наличие верхней оценки для решений уравнений (2.2″), (3.2″). Для линейных задач такие оценки давались в работах [41, 58, 82, 206, 222, 226]. Приведем, например, результат из [222], касающийся рассматриваемого случая линейных уравнений в гильбертовых пространствах.

Лемма 3. *Пусть уравнение* (1.1) *разрешимо и* $0 \notin \bar{Z}$. *Если при* $0 \leqslant \sigma \leqslant \sigma_0$ *выполнено неравенство* $\| u_\sigma \| / \sigma > C > 1$ ($C = \text{const}$), *то для решения* $\bar{\alpha}_\delta$ *уравнения* (3.2″) *имеет место оценка*

$$\bar{\alpha}_\delta < \bar{\alpha},$$

$$\bar{\alpha} = \| A_h \| \cdot \left(h + \frac{\sqrt{h^2 + (\tilde{\alpha} + h^2)(C^2 - 1)} + h}{C^2 - 1} \right),$$

где

$$\tilde{\alpha} = \| A_h \|^2 C\sigma / (\| u_\sigma \| - C\sigma).$$

Результат леммы 2 и результаты, аналогичные лемме 3, позволяют применить для приближенного нахождения корней уравнений (2.2″), (3.2″) комбинированный метод хорд и касательных [21] на начальных итерациях и метод Ньютона на последующих.

§ 5. О связи алгоритмов обобщенных принципов с некоторыми другими алгоритмами решения операторных уравнений

В § 1.12 рассматривалась связь обобщенных принципов невязки и квазирешений для экстремальных задач с обобщенным методом невязки и методом квазирешений. Эти результаты могут быть конкретизированы для задачи отыскания Ω-оптимальных квазирешений операторных уравнений (1.1) следующим образом.

Обобщенный метод невязки (о.м.н.) для задачи (1.1) заключается в том, что в качестве ее приближенного решения принимается какой-либо элемент $z_\delta \in D$, для которого

$$\Omega(z_\delta) = \inf\{\Omega(z)\colon z \in D,\ \ \rho(A_h z, u_\sigma) \leqslant \lambda_\delta + \sigma + \psi(h, \Omega(z))\} \equiv \tilde{\Omega}_\delta, \qquad (1)$$

если такой элемент существует.

Справедливы следующие аналоги теорем 1.12.1—1.12.3, устанавливающие свойства о.м.н. для операторных уравнений и его связь с о.п.н.

Теорема 1. *Пусть выполнены предположения* **1—3** *из* § 1 *и, кроме того, функционал* $\Omega(z)$ τ-*секвенциально* **непрерывен на** D. *Тогда задача* (1) *разрешима.*

Теорема 2. *Предположим, что выполнены предположения* 1—4 *из* § 1 *и уравнение* (2.2) *для нахождения параметра регуляризации по алгоритму о.п.н. имеет обычное решение* $\alpha_\delta > 0$. *Тогда задача* (1) *разрешима и любое приближенное решение* z^{α_δ}, *полученное по алгоритму о.п.н., является приближенным решением, полученным по о.м.н. (т. е. решением задачи* (1)).

Теорема 3. *Пусть выполнены предположения* 1—4 *из* § 1 *и, кроме того, числа* ρ_0, ρ_∞, *определенные в* § 2, *таковы, что* $\rho_0 < 0$, $\rho_\infty > 0$ *для рассматриваемого* $\delta = (h, \sigma)$. *Пусть также задача* (1.9) *минимизации сглаживающего функционала имеет при каждом* $\alpha > 0$ *единственное решение* z^α, *причем для любой пары различных значений* α_1, $\alpha_2 > 0$ *реализуется требование* $z^{\alpha_1} \neq z^{\alpha_2}$. *Тогда единственное приближенное решение* z^{α_δ}, *выбираемое по о.п.н. (см. теорему* 2.2), *является единственным решением задачи* (1).

Теорема 3 устанавливает эквивалентность о.п.н. и о.м.н. в случае, указанном в условиях. Если условия теоремы 3 нарушаются, то вопрос об эквивалентности о.п.н. и о.м.н. решается, вообще говоря, отрицательно. Для того чтобы убедиться в этом, достаточно рассмотреть пример 2.1, в котором требование единственности решения задачи (1.9) нарушается при $\alpha = \alpha_0$. В этом примере показано, что алгоритм о.п.н дает приближенное решение $z^{\alpha_\delta} = z_0$. На рис. 5 показано множество элементов z, подчиненных ограничению обобщенного метода невязки: $\rho(A_h, u_\sigma) \leqslant \lambda_\delta + \sigma + \psi(h, \Omega(z))$. Ясно, что о.м.н. дает в качестве приближенного решения такой элемент z_δ, для которого выполнено неравенство $\|z_\delta\| < \|z_0\| = \|z^{\alpha_\delta}\|$. Таким образом, в данном примере алгоритмы о.п.н. и о.м.н. не эквивалентны. Другие примеры неэквивалентности о.п.н. и о.м.н. приведены в [88], где рассматривается случай линейных операторных уравнений.

Вопрос о сходимости приближений, полученных по алгоритму о.м.н., решает

Теорема 4. *Предположим, что для произвольной последовательности* $\{\varepsilon_n\}$, *где* $\varepsilon_n > 0$, *сходящейся к нулю при* $n \to \infty$, *а также для любой последовательности* $\{\delta_n\}$, *также сходящейся к нулю при* $n \to \infty$, *найдены элементы* z_n, *«почти удовлетворяющие обобщенному методу невязки», т. е. такие* z_n, *для которых выполнены ограничения задачи* (1) *и неравенство* $\tilde{\Omega}_{\delta_n} \leqslant \Omega(z_n) \leqslant \tilde{\Omega}_{\delta_n} + \varepsilon_n$. *Тогда* $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n \to \infty$. *Если задача* (1) *разрешима и для каждого* δ_n *найдено какое-либо ее решение* z_n, *то также* $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n \to \infty$.

В случае, когда имеется априорная информация о разрешимости уравнения (1.1) на множестве D, возможна иная (модифицированная) форма алгоритма о.м.н., не использующая величин λ_δ. При этом приближенное решение z_δ находится как решение экстремальной задачи: найти элементы $z_\delta \in D$, для которых

$$\Omega(z_\delta) = \inf\{\Omega(z)\colon z \in D,\ \ \rho(A_h z, u_\sigma) \leqslant \sigma + \psi(h, \Omega(z))\}. \qquad (2)$$

Эта форма обобщенного метода невязки связана с модифицированным о.п.н. (см. § 3). Вопросы разрешимости задачи (2), ее связь с м.о.п.н., сходимость приближений z_δ к множеству Ω-оптимальных решений уравнения (1.1) исследуются так же, как и в теоремах из §1.12 и в их аналогах — теоремах 1—4. При этом фигурирующие в доказательствах числа $J^*=\mu_0$, λ_δ заменяются нулями.

Метод квазирешений (м.к.) для операторных уравнений [84—88] использует дополнительную информацию о квазирешениях задачи (1.1): считается, что некоторая оценка ω_δ для величины $\bar{\Omega}$ известна. Тогда в качестве приближений к множеству $\bar{Z}$ предлагается брать решения экстремальной задачи: найти элементы $z_\delta \in D$ такие, что справедливо равенство

$$\rho(A_h z_\delta, u_\sigma) = \inf\{\rho(A_h z, u_\sigma):\ z\in D,\ \ \Omega(z)\leqslant\omega_\delta\}. \tag{3}$$

Эта задача, как легко видеть, является частным случаем экстремальной задачи (1.12.10) — метода квазирешений для экстремальных задач — и поэтому обладает всеми свойствами, обоснованными в теоремах 1.12.5 — 1.12.7.

Теорема 5. *Пусть выполнены предположения* 1—3 *из* § 1 *и число* ω_δ *таково, что* $\omega_\delta \geqslant \Omega^*$. *Тогда задача* (3) *разрешима.*

Теорема 6. *Если выполнены предположения* 1—3 *из* § 1, *а число* ω_δ *взято так, что при каждом* δ $\omega_\delta \geqslant \bar{\Omega}$ *и* $\omega_\delta \to \bar{\Omega}$ *при* $\|\delta\| \to 0$, *то для любой последовательности* $\{\delta_n\}$, *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательность* $\{z_n\}$, *где* $z_n \equiv z_{\delta_n}$, *решений задачи* (3) *обладает следующими свойствами*: $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n\to\infty$.

Теорема 7. *Предположим, что выполнены предположения* 1—4 *из* § 1, *а для числа* ω_δ *справедливо неравенство* $\Omega^* < \omega_\delta < \Omega_\delta \equiv \gamma(+0)$. *Пусть, кроме того, задача* (1.9) *минимизации сглаживающего функционала имеет при каждом* $\alpha>0$ *единственное решение* z^α *такое, что* $z^{\alpha_1} \neq z^{\alpha_2}$ *для любых* $\alpha_1, \alpha_2 > 0$ $(\alpha_1 \neq \alpha_2)$. *Тогда единственное приближенное решение* z^{α_δ}, *определяемое по алгоритму о.п.к.* (*см. теорему* 2.5), *является единственным решением задачи* (3).

Эта теорема устанавливает эквивалентность о.п.к. и м.к. при выполнении указанных условий. В общем случае алгоритмы о.п.к. и м.к. не являются эквивалентными. Чтобы убедиться в этом, рассмотрим пример 3.1.

Зададим величину ω_δ так, чтобы $\omega_\delta > \bar{\Omega}$ (в примере 3.1 имеем $\bar{\Omega}=1$). Найдем приближенное решение задачи из примера 3.1 по алгоритму о.п.к. Для этого используем выражение для экстремали z^α, найденное в этом примере:

$$\Omega(z^\alpha)=\gamma(\alpha)=\{1+h^{-1/2}\colon 0<\alpha\leqslant h^{3/2};\ \ 1\colon h^{3/2}\leqslant\alpha\leqslant 1;\ \ 0\colon \alpha\geqslant 1\}.$$

График этой функции приведен на рис. 13. Ясно, что $\alpha_\delta = h^{3/2}$. Найдем теперь приближенное решение z^{α_δ} по правилу отбора 2.2. Для этого заметим, что в данном примере функция

$$\Phi(z) \equiv \|A_h z - u_\sigma\| + h\Omega(z) = |A_h z - \bar{u}| + h\,|z+1|,$$

используемая для вычисления обобщенной меры несовместности λ_δ, является кусочно-линейной. Поэтому

$$\lambda_\delta = \inf \Phi(z) = \min\{\Phi(-1),\ \Phi(0),\ \Phi(h^{-1/2})\} = \\ = \min\{2+h,\ 1+2h,\ 2+h\} = 1+2h$$

при $0<h<1$. Обращаясь к правилу 2.2, можно заметить, что выполнено неравенство (2.4):

$$\beta(\alpha_2)=\beta(qh^{3/2})=1+h<C\Pi_\delta(z^{\alpha_1})-(C-1)\lambda_\delta=$$
$$=C\left[h\,|z^{\alpha_\delta/q}+1|+\lambda_\delta\right]-(C-1)\lambda_\delta=Ch(h^{-1/2}+1)+1+2h.$$

Тогда, согласно правилу отбора, необходимо выбрать в качестве приближенного решения такой элемент z^{α_δ} из множества $Z^{\alpha_\delta}=\{h^{-1/2}, 0\}$, для которого выполнено неравенство $\Omega(z^{\alpha_\delta})\leqslant\omega_\delta$, т. е. $|1+z^{\alpha_\delta}|\leqslant\omega_\delta$. Таким элементом множества Z^{α_δ} будет $z^{\alpha_\delta}=0=\bar{z}$.

С другой стороны, по методу квазирешений необходимо минимизировать невязку $N_\delta(z)$, график которой представлен на рис. 11 на множестве $\{z\colon \Omega(z)\leqslant\omega_\delta\}=\{z\colon -1-\omega_\delta\leqslant z\leqslant -1+\omega_\delta\}$. Решением этой задачи будет элемент $z_\delta=\omega_\delta-1>0$. Таким образом, $z_\delta\neq z^{\alpha_\delta}$, и в данном примере алгоритмы о.п.к. и м.к. не будут эквивалентными.

Обратимся теперь к рассмотрению связи алгоритмов о.п.н. и о.п.с.ф. с так называемым альтернативным принципом выбора параметра регуляризации. Альтернативный принцип был предложен в работах [63, 65] для решения нелинейных операторных уравнений с точно заданным оператором ($h=0$) и в дальнейшем изучался в работах [116, 206].

Будем предполагать, что выполнены следующие условия. Пусть Z — банахово пространство, D — замкнутое множество в Z, U — метрическое пространство с метрикой ρ. Будем считать, что рассматривается уравнение вида (1.1) с точно заданным, непрерывным и взаимно однозначным оператором A, действующим из D в U. При этом считаем, что для правой части $\bar{u}\in U$ этого уравнения существует единственное решение $\bar{z}\in D$, которое истокообразно представимо с помощью линейного, вполне непрерывного, взаимно однозначного оператора B, действующего из рефлексивного банахова пространства V в Z: $\bar{z}=B\bar{v}$ ($\bar{v}\in V$). Далее предположим, что известен элемент $u_\sigma\in U$, представляющий собой приближение к $\bar{u}$, удовлетворяющее условию $\rho(\bar{u}, u_\sigma)<\sigma$, причем число σ — оценка погрешности правой части уравнения (1.1) — известно. Таким образом, в рассматриваемой задаче $\delta=(0, \sigma)$.

Введем сглаживающий функционал вида

$$M^\alpha[v]=\alpha g(\|v\|)+f\left[\rho(ABv, u_\sigma)\right]\equiv\alpha\Omega(v)+I_\delta(v).$$

В этом функционале вспомогательная функция $g(x)$ определена при $x\geqslant 0$, непрерывна, неотрицательна, монотонно возрастает и удовлетворяет требованию $g(+\infty)=+\infty$. Другая вспомогательная функция $f(t)$ задана при $t\geqslant 0$ и обладает аналогичными свойствами. Из общих результатов теоремы 1.4.2, условий § 2.1 и требований постановки рассматриваемой задачи ясно, что задача минимизации сглаживающего функционала, определенного выше, на множестве $\mathcal{V}=\{v\in V\colon Bv\in D\}\ni\bar{v}$ имеет при каждом $\alpha>0$ непустое множество решений V^α. Используя элементы v^α этих множеств V^α, можно по аналогии с § 2 определить при $\alpha>0$ следующие, вообще говоря, многозначные функции:

$$\beta(\alpha)\equiv f[\rho(ABv^\alpha, u_\sigma)]\equiv I_\delta(v^\alpha),$$
$$\gamma(\alpha)\equiv g(\|v^\alpha\|)=\Omega(v^\alpha),$$

а также однозначную функцию $\varphi(\alpha)\equiv M^\alpha[v^\alpha]$. Из общих результатов полученных в § 1.6, выводятся следующие свойства этих функций.

1. Функции $\beta(\alpha)$, $\gamma(\alpha)$ непрерывны и однозначны при $\alpha>0$ всюду, за исключением, быть может, не более чем счетного числа их общих точек многозначности, являющихся одновременно точками разрыва первого рода.

2. Функция $\beta(\alpha)$ монотонно не убывает, а $\gamma(\alpha)$ монотонно не возрастает при $\alpha>0$.

3. Функция $\varphi(\alpha)$ непрерывна. Она монотонно возрастает при $\alpha>0$ в случае, когда $\Omega^*=\inf\{g(\|v\|): v\in\mathscr{V}\}>0$; если $\Omega^*=0$, то $\varphi(\alpha)$ монотонно возрастает при $0<\alpha<\alpha_{\min}$ и постоянна при $\alpha\geqslant\alpha_{\min}$. Здесь $\alpha_{\min}$ — наименьшее из решений (если они существуют) уравнения с монотонной функцией: $\gamma(\alpha)=0$.

4. Существуют пределы

$$\lim_{\alpha\to+0}\beta(\alpha)=\lim_{\alpha\to+0}\varphi(\alpha)=f(\mu_\delta),$$

$$\lim_{\alpha\to+\infty}\beta(\alpha)=\lim_{\alpha\to+\infty}\varphi(\alpha)=\nu_\delta\equiv\inf\{f[\rho(ABv, u_\sigma)]: v\in\mathscr{V}_0\},$$

$$\mathscr{V}_0\equiv\{v\in\mathscr{V}: \Omega(v)=\Omega^*\},$$

$$\lim_{\alpha\to+\infty}\gamma(\alpha)=\Omega^*\equiv\inf\{\Omega(v): v\in\mathscr{V}\}\geqslant 0.$$

Справедлива

Лемма 1. *Пусть кроме условий постановки задачи выполнено требование* $\bar v\notin\mathscr{V}_0$. *Тогда существует такое число* $\sigma_0>0$, *что для всех* σ $(0<\sigma<\sigma_0)$ *выполнено неравенство* $\nu_\delta>f(\sigma_0)>f(\sigma)$.

Чтобы убедиться в этом, обратимся к доказательству теоремы 1.7.3. Там установлено неравенство (1.7.4), которое с учетом разрешимости рассматриваемого операторного уравнения принимает вид

$$\nu_0=\inf\{f[\rho(ABv, \bar u)]: v\in\mathscr{V}_0\}>\inf\{f[\rho(ABv, \bar u)]: v\in\mathscr{V}\}=$$
$$=f[\rho(AB\bar v, \bar u)]=f(0).$$

Отсюда следует, что $f^{-1}(\nu_0)>0$. Выберем теперь $\sigma_0\equiv f^{-1}(\nu_0)/2$. Тогда из оценки (1.7.5) с учетом равенства $\Psi(\delta, \Omega)\equiv\sigma$, а также с использованием монотонности функции f получим неравенство, справедливое для всех σ $(0<\sigma<\sigma_0)$:

$$\nu_\delta\geqslant f[f^{-1}(\nu_0)-\sigma]>f[f^{-1}(\nu_0)-\sigma_0]=f(\sigma_0)>f(\sigma).$$

Лемма доказана.

В дальнейшем будем считать, что условие $\bar v\notin\mathscr{V}_0$ леммы 1 выполнено.

Определим теперь оператор выбора Q, который сопоставляет каждому множеству V^α некоторый его элемент v^α. Используем выбранные таким образом элементы v^α для вычисления функций $\beta(\alpha)$, $\gamma(\alpha)$, $\varphi(\alpha)$. Результатом этих вычислений окажутся соответственно уже однозначные функции $\beta_Q(\alpha)$, $\gamma_Q(\alpha)$, $\varphi_Q(\alpha)$. Ясно, что $\varphi_Q(\alpha)=\varphi(\alpha)$ для любого $\alpha>0$, а $\beta_Q(\alpha)=\beta(\alpha)$, $\gamma_Q(\alpha)=\gamma(\alpha)$ в точках непрерывности этих функций. Функции $\beta_Q(\alpha)$, $\gamma_Q(\alpha)$, $\varphi_Q(\alpha)$ обладают свойствами 1—4, перечисленными выше.

С помощью этих функций зададим при данном $\delta=(0, \sigma)$ множества значений параметра регуляризации по следующим правилам:

$$\begin{aligned}
&\mathfrak{A}_1=\mathfrak{A}_1(\delta)=\{\alpha\colon \alpha>0,\ \beta_Q(\alpha)\geqslant f(\sigma)\},\\
&\mathfrak{A}_2=\mathfrak{A}_2(\delta)=\{\alpha\colon \alpha>0,\ \beta_Q(\alpha)\leqslant Cf(\sigma)\},\\
&\mathfrak{B}_1=\mathfrak{B}_1(\delta)=\{\alpha\colon \alpha>0,\ \beta_Q(\alpha)\leqslant f(\sigma)\},\\
&\mathfrak{B}_2=\mathfrak{B}_2(\delta)=\{\alpha\colon \alpha>0,\ \varphi_Q(\alpha)\geqslant Cf(\sigma)\}.
\end{aligned} \tag{4}$$

Используемый здесь параметр C не зависит от σ и удовлетворяет при всех σ $(0<\sigma<\sigma_0)$ неравенству $1<C<\nu_\delta/f(\sigma_0)$.

Возможность выбора такого числа C следует из леммы 1.

Лемма 2. *Множества* (4) *непусты при* $0<\sigma<\sigma_0$.

Доказательство. Из свойства 4 функций $\beta_Q(\alpha)$, $\varphi_Q(\alpha)$ с учетом условия выбора константы C получим

$$\lim_{\alpha\to+\infty}\beta_Q(\alpha)=\nu_\delta>Cf(\sigma_0)>Cf(\sigma)>f(\sigma),$$

$$\lim_{\alpha\to+0}\beta_Q(\alpha)=f(\mu_\delta)=\inf\{f[\rho(ABv, u_\sigma)]\colon v\in\mathscr{V}\}\leqslant$$
$$\leqslant f[\rho(AB\bar{v}, u_\sigma)]=f[\rho(\bar{u}, u_\sigma)]<f(\sigma)<Cf(\sigma),$$

$$\lim_{\alpha\to+0}\beta_Q(\alpha)=f(\mu_\delta)<f(\sigma), \quad \lim_{\alpha\to+\infty}\varphi_Q(\alpha)=\nu_\delta>Cf(\sigma).$$

Из этих соотношений вытекает непустота множеств (4).

Лемма 3. *Множество* $\mathscr{E}=\mathscr{E}(\delta)=(\mathfrak{A}_1\cap\mathfrak{A}_2)\cup(\mathfrak{B}_1\cap\mathfrak{B}_2)$ *непусто.*

Доказательство. Предположим, что $\mathfrak{A}_1\cap\mathfrak{A}_2=\varnothing$. Это означает, что выбор α из неравенства $f(\sigma)\leqslant\beta_Q(\alpha)\leqslant Cf(\sigma)$ невозможен. Покажем, что в этом случае $\mathfrak{B}_1\cap\mathfrak{B}_2\neq\varnothing$.

В силу монотонного неубывания функции $\beta_Q(\alpha)$ и вследствие отмеченного в лемме 2 неравенства $\beta_Q(+\infty)>f(\sigma)$ найдется число

$$\alpha_0=\sup\{\alpha\colon \alpha>0,\ \beta_Q(\alpha)\leqslant f(\sigma)\}>0.$$

Рассмотрим последовательность $\{\alpha_n\}$, обладающую тем свойством, что $\alpha_n\to\alpha_0+0$ при $n\to\infty$ $(\alpha_n>\alpha_0)$. Тогда вследствие определения числа α_0 и по предположению о пустоте множества $\mathfrak{A}_1\cap\mathfrak{A}_2$ получим неравенства

$$\beta_Q(\alpha_n)>f(\sigma), \quad \beta_Q(\alpha_n)\geqslant Cf(\sigma).$$

Из них следует, что

$$\varphi(\alpha_n)=\beta_Q(\alpha_n)+\alpha_n\gamma_Q(\alpha_n)\geqslant Cf(\sigma)+\alpha_n\gamma_Q(\alpha_n).$$

Переходя в последнем соотношении к пределу при $n\to\infty$ и учитывая непрерывность функции $\varphi(\alpha)$, получаем

$$\varphi(\alpha_0)=\varphi(\alpha_0+0)\geqslant Cf(\sigma)+\alpha_0\gamma_Q(\alpha)(\alpha_0+0).$$

Возможны два случая. В первом $\gamma_Q(\alpha_0+0)>0$. Тогда $\varphi(\alpha_0)>Cf(\sigma)$, и в силу непрерывности функции $\varphi(\alpha)$ найдется такая окрестность $O(\alpha_0)$ точки $\alpha_0>0$, где выполнено неравенство $\varphi(\alpha)\geqslant Cf(\sigma)$. Выбрав из этой окрестности число $\alpha^*<\alpha_0$, найдем по определению числа α_0 и окрестности $O(\alpha_0)$, что

$$\varphi(\alpha^*)\geqslant Cf(\sigma),\quad \beta_Q(\alpha^*)\leqslant f(\sigma).$$

Таким образом, $\alpha^*\in\mathfrak{B}_1\cap\mathfrak{B}_2$.

Во втором случае $\gamma_Q(\alpha_0+0)=0$. Это возможно лишь в случае, если $\Omega^*=0$. Тогда по свойству 3 функции $\varphi(\alpha)$, отмеченному выше, $\varphi(\alpha)=\nu_\delta>Cf(\sigma)$ при любом $\alpha\geqslant\alpha_0$; в частности, $\varphi(\alpha_0)>Cf(\sigma)$. Дальнейшие рассуждения проводятся, как и при рассмотрении первого случая. Лемма доказана.

Сформулируем альтернативный принцип выбора параметра регуляризации: в качестве значения α_δ выбирается любой элемент множества $\mathscr{E}$, а в качестве приближенного решения задачи — элемент $z_\delta=BQV^{\alpha_\delta}\equiv Bv^{\alpha_\delta}$.

Теорема 8. *Если элементы v^{α_δ}, $z_\delta=Bv^{\alpha_\delta}$ найдены по альтернативному принципу при $\|\delta\|\leqslant\Delta_0$, то при $\|\delta\|\to 0$ элементы v^{α_δ} слабо сходятся к $\bar v$, а элементы z_δ сильно сходятся к $\bar z=B\bar v$.*

Доказательство. Предположим, что $\alpha_n\equiv\alpha_{\delta_n}\in\mathfrak{A}_1(\delta_n)\cap\mathfrak{A}_2(\delta_n)$. Тогда в силу экстремальности элемента $v_n\equiv v^{\alpha_n}$

$$f(\sigma_n)+\alpha_n g(\|v_n\|)\leqslant f[\rho(ABv_n,u_{\sigma_n})]+\alpha_n g(\|v_n\|)\leqslant$$
$$\leqslant f[\rho(AB\bar v,u_{\sigma_n})]+\alpha_n g(\|\bar v\|)\leqslant f(\sigma_n)+\alpha_n g(\|\bar v\|),$$

откуда ясно, что $g(\|v_n\|)\leqslant g(\|\bar v\|)$. Если же $\alpha_n\in\mathfrak{B}_1(\delta_n)\cap\mathfrak{B}_2(\delta_n)$, то

$$\varphi(\alpha_n)\geqslant Cf(\sigma_n),$$
$$\varphi(\alpha_n)\leqslant f[\rho(AB\bar v,u_{\sigma_n})]+\alpha_n g(\|\bar v\|)\leqslant f(\sigma_n)+\alpha_n g(\|\bar v\|).$$

Из этих неравенств получим

$$\alpha_n\geqslant(C-1)f(\sigma_n)/g(\|\bar v\|),$$

$$g(\|v_n\|)\leqslant\varphi(\alpha_n)/\alpha_n\leqslant[f(\sigma_n)+\alpha_n g(\|\bar v\|)]/\alpha_n\leqslant$$
$$\leqslant g(\|\bar v\|)/(C-1)+g(\|\bar v\|)=Cg(\|\bar v\|)/(C-1).$$

Таким образом, для элементов v_n справедлива оценка

$$\|v_n\|\leqslant\max\left\{g^{-1}\left[\frac{Cg(\|\bar v\|)}{C-1}\right],\ \|\bar v\|\right\}=g^{-1}\left[\frac{Cg(\|\bar v\|)}{C-1}\right],$$

а также вытекающее из принадлежности $\alpha_n\in\mathscr{E}(\delta_n)$ неравенство

$$I_{\delta_n}(v_n)=\beta_Q(\alpha_n)\leqslant Cf(\sigma_n)\to 0,\quad n\to\infty.$$

Эти соотношения вместе с предположением единственности решения $t=B\bar{v}$ на основании леммы 1.13.2 доказывают рассматриваемую теорему.

Отметим, что альтернативный принцип гарантирует сходимость приближений независимо от вида оператора выбора Q.

Для установления связи альтернативного принципа и алгоритма о.п.н. рассмотрим вариант модифицированного обобщенного принципа невязки для исследуемого уравнения (1.1) с точным оператором. В этом варианте положим $K=C$ (см. (3.1)). Тогда параметр регуляризации, определяемый по м.о.п.н., находится из уравнения

$$\beta(\alpha)-Cf(\sigma)=0 \tag{5}$$

(см. (3.1), (3.2)). Это уравнение имеет при $\sigma>0$ решение $\alpha_\sigma>0$. Это следует из неравенств $\beta_\infty\equiv\nu_\delta-Cf(\sigma)>0$, $\beta_0\equiv f(\mu_\delta)-Cf(\sigma)<0$, полученных в лемме 2, и из леммы 1.6.7. Кроме того, из монотонности функции $\beta(\alpha)$ и следствия 1.6.5 получаются неравенства

$$\beta(\alpha_2)\geqslant\beta(\alpha_\sigma+0)\geqslant Cf(\sigma)\equiv\bar{\pi}(\alpha_1).$$

Здесь числа $\alpha_{1,2}$ определяются согласно алгоритму о.п.н.: $\alpha_1=\alpha_\sigma/q$, $\alpha_2=\alpha_\sigma q$ $(q=\mathrm{const}>1)$. Тем самым для рассматриваемого варианта алгоритма о.п.н. всегда реализуется требование (3.3) первой части правила отбора 3.1. Поэтому приближенное решение задачи выбирается так, чтобы выполнялось условие (3.5), которое в данном случае принимает вид

$$\beta(\alpha_\sigma)=I_\delta(v^{\alpha_\sigma})=f\left[\rho(ABv^{\alpha_\sigma}, u_\sigma)\right]\leqslant Cf(\sigma). \tag{6}$$

В частности, можно принять в качестве приближенного решения рассматриваемого операторного уравнения элемент $z_\delta=Bv_-^{\alpha_\sigma}$. Если при этом определить оператор Q равенством $Q\bar{V}^\alpha=v_-^\alpha$ для любого $\alpha>0$, то в соответствии с (5), (6) выбор α_σ как решения уравнения с монотонной функцией

$$\beta_Q(\alpha)=Cf(\sigma) \tag{7}$$

порождает приближенное решение $z_\delta=BQV^{\alpha_\sigma}=Bv_-^{\alpha_\sigma}$ рассматриваемой задачи. Для этого приближения справедлива теорема сходимости 3.1. Найденный таким методом параметр $\alpha_\sigma>0$ обладает следующим свойством.

Теорема 9. *Справедливо включение* $\alpha_\sigma\in\mathscr{E}(\delta)$ *для любого* σ $(0<\sigma<\sigma_0)$.

Действительно, если $\beta_Q(\alpha_\sigma)\geqslant f(\sigma)$, то $\alpha_\sigma\in\mathfrak{A}_1$, а вследствие неравенства (6) имеем $\alpha_\sigma\in\mathfrak{A}_2$. Поэтому $\alpha_\sigma\in\mathfrak{A}_1\cap\mathfrak{A}_2\subset\mathscr{E}$. Если же $\beta_Q(\alpha_\sigma)\leqslant f(\sigma)$, то $\alpha_\sigma\in\mathfrak{B}_1$. Тогда из следствия 1.6.5 для решения α_σ уравнения (5) (или (7)) получим с учетом непрерывности функции $\varphi(\alpha)$

$$\varphi(\alpha_\sigma)=\alpha_\sigma\gamma(\alpha_\sigma+0)+\beta(\alpha_\sigma+0)\geqslant\beta(\alpha_\sigma+0)=\beta_Q(\alpha_\sigma+0)\geqslant Cf(\sigma),$$

т. е. $\alpha_\sigma\in\mathfrak{B}_2$. Итак, $\alpha_\sigma\in\mathfrak{B}_1\cap\mathfrak{B}_2\subset\mathscr{E}$. Теорема доказана.

Теорема 9 показывает, что параметр регуляризации α_σ, найденный по рассмотренному варианту алгоритма м.о.п.н., удовлетворяет альтернативному принципу выбора параметра.

Рассмотрим теперь вариант *принципа сглаживающего функционала* (п.с.ф), который изучался в работах [142, 150, 151]. В этом варианте параметр α_σ выбирается как обычное решение уравнения с монотонной и непрерывной функцией:

$$\varphi(\alpha)=Cf(\sigma).$$

В качестве приближенного решения рассматриваемого операторного уравнения берется элемент $z_\delta=Bv^{\alpha_\sigma}$, отвечающий любому $v^{\alpha_\sigma}\in V^{\alpha_\sigma}$, Как явствует из свойства возрастания функции $\varphi(\alpha)$, леммы 1.6.7 и доказанных в лемме 2 соотношений $\varphi(+0)<Cf(\sigma)$, $\varphi(+\infty)>$ $>Cf(\sigma)$, уравнение (8) имеет при каждом σ $(0<\sigma<\sigma_0)$ единственное решение $\alpha_\sigma>0$.

Теорема 10. *Для произвольного оператора Q число α_σ, выбранное как решение уравнения* (8), *удовлетворяет альтернативному принципу*

Чтобы убедиться в этом, достаточно заметить, что из (8) вытекают неравенства

$$\beta_Q(\alpha_\sigma)\leqslant\varphi(\alpha_\sigma)=Cf(\sigma),$$

согласно которым $\alpha_\sigma\in\mathfrak{A}_2$, $\alpha_\sigma\in\mathfrak{B}_2$. Поскольку для α_σ обязательно выполнено одно из включений $\alpha_\sigma\in\mathfrak{A}_1$ или $\alpha_\sigma\in\mathfrak{B}_1$, то $\alpha_\sigma\in\mathcal{E}$.

Таким образом, алгоритмы м.о.п.н. и п.с.ф. связаны с альтернативным принципом выбора параметра регуляризации. Этот принцип дает широкий спектр допустимых значений параметра α и не требует для своей реализации специального отбора экстремали. Однако теория альтернативного принципа развита лишь для случая точного задания оператора.

§ 6. Оптимальность по порядку точности алгоритмов обобщенных принципов при решении операторных уравнений

В § 1, 2 отмечено, что алгоритмы о.п.н., о.п.к., о.п.с.ф. для решения некорректных экстремальных задач могут быть использованы при решении вариационным методом операторных уравнений. При этом для них, естественно, остаются справедливым все свойства, установленные в гл. 1. Это касается, в частности, и результатов об оптимальном порядке точности алгоритмов о.п.н., о.п.к., о.п.с.ф. Сформулируем требования, при выполнении которых результаты § 1.13 переносятся и на случай решения операторных уравнений.

Пусть Z — банахово пространство, D — замкнутое множество, принадлежащее Z, а U — метрическое пространство с топологией, порождаемой метрикой ρ. Пусть также A_0 — непрерывный, взаимно однозначный оператор, действующий из D в U, а $\bar{u}_0\in U$ — некоторый элемент. Будем предполагать, что уравнение

$$A_0z=\bar{u}_0,\quad z\in D,$$

имеет единственное квазирешение $\bar{z}_0$ на множестве D, причем известно, что это квазирешение истокообразно представимо с помощью оператора B: $\bar{z}_0=B\bar{v}$ $(\bar{v}\in V,\ \bar{v}\neq0)$. Здесь B — линейный, вполне непрерывный, взаимно однозначный оператор, действующий из рефлексивного банахова пространства V в Z. Далее предположим, что вместо неизвестных точных данных $(A_0, \bar{u}_0)$ известны их приближения (A_h, u_σ). При этом A_h — непрерывный, взаимно однозначный оператор, действующий из D в U, — удовлетворяет условию аппроксимации

$$\rho(A_h z, A_0 z) \leqslant \psi(h, \|z\|) \quad \forall z \in D,$$

а для $u_\sigma \in U$ выполнено неравенство $\rho(\bar{u}_0, u_\sigma) \leqslant \sigma$. Мера аппроксимации ψ известна и удовлетворяет требованиям из § 1, 2; неотрицательные числа σ, h, определяющие погрешности приближенных данных, также считаются известными.

Вводя функционалы невязки

$$J_0(z) = \rho(A_0 z, \bar{u}_0), \quad J_\delta(z) = \rho(A_h z, u_\sigma), \quad \delta \equiv (h, \sigma),$$

меру аппроксимации $\psi_0(h, \|z\|) = \sigma + \psi(h, \|z\|)$ и регуляризующий функционал вида $\Omega(v) = g(\|v\|)$, где функция $g(x)$ определена при $x \geqslant 0$, непрерывна, неотрицательна, монотонно возрастает и подчиняется условию $g(+\infty) = +\infty$, можно убедиться, что выполнены все условия из § 1.13 для решения задачи о поиске Ω-оптимального решения вариационной задачи вида (1.13.1) с указанным функционалом J_0. Таким образом, для приближенного устойчивого нахождения квазирешения $\bar{z}_0$ могут быть использованы варианты алгоритмов о.п.н., о.п.к., о.п.с.ф., изученные в § 1.13. Для полученных с их помощью приближенных решений $z_\delta = Bv^{\alpha_\delta}$ остаются справедливыми результаты теорем 1.13.1—1.13.7 и, в частности, результат теоремы 1.13.8 об оптимальном порядке точности этих алгоритмов. Такими же свойствами оптимальности будут обладать приближенные решения задачи (1), полученные с помощью вариантов метода невязки или метода квазирешений, который использован в § 1.13.

Рассмотрим теперь важный частный случай задачи (1), когда имеется дополнительная априорная информация о разрешимости операторного уравнения [133, 136]. Как указано в § 3, в этом случае $\mu_0 = 0$ и величину обобщенной меры несовместности λ_δ можно исключить из алгоритмов. При этом форма алгоритмов слегка меняется (см. § 3). Рассмотрим эти модификации алгоритмов в данной постановке задачи.

Пусть $\mathscr{V} \equiv \{v \in V\colon\ Bv \in D\}$. Определим на $\mathscr{V}$ сглаживающий функционал

$$M^\alpha[v] = \alpha g(\|v\|) + f[\rho(A_h Bv, u_\sigma)] \equiv \alpha\Omega(v) + I_\delta(v), \tag{2}$$

в котором $g(x)$ удовлетворяет перечисленным выше условиям, а функция $f(t)$, заданная при $t \geqslant 0$, обладает теми же свойствами, что и $g(x)$, и, кроме того, является выпуклой (вниз). Введем задачу минимизации сглаживающего функционала: найти элементы $v^\alpha \in \mathscr{V}$, для которых

$$M^\alpha[v^\alpha] = \inf\{M^\alpha[v]\colon\ v \in \mathscr{V}\}.$$

Из условий постановки задачи данного параграфа, а также из § 1, 1.13 ясно, что такая экстремальная задача имеет при каждом $\alpha > 0$ непустое множество решений V^α, с помощью элементов которого можно определить многозначные функции, аналогичные функциям (3.1), но несколько отличающиеся от функций (1.13.8):

$$\beta(\alpha) \equiv f\left[\rho(A_h Bv^\alpha, u_\sigma)\right] = I_\delta(v^\alpha), \quad \gamma(\alpha) = \Omega(v^\alpha),$$
$$\bar{\rho}(\alpha) \equiv I_\delta(v^\alpha) - \bar{\Pi}_\delta(v^\alpha), \quad \bar{\Pi}_\delta(v^\alpha) \equiv f\left[\Psi(\delta, \Omega(v^\alpha))\right],$$
$$\varepsilon(\alpha) \equiv \varphi(\alpha) - f\left[C\Psi(\delta, \Omega(v^\alpha))\right], \quad C = \mathrm{const} > 1,$$
$$\varphi(\alpha) \equiv M^\alpha\left[v^\alpha\right] \quad \forall v^\alpha \in V^\alpha, \quad \Psi(\delta, \Omega(v)) \equiv \sigma + \psi(h, \|B\| \cdot \|v\|).$$

Для этих функций справедливы стандартные свойства, перечисленные в § 1.6—1.9.

Алгоритм *модифицированного обобщенного принципа невязки* (м.о.п.н.) для решения уравнения (1) при наличии априорной информации о его разрешимости на множестве D и об истокообразной представимости решения с помощью оператора B по структуре своей близок к введенному в § 3. Однако правило отбора элемента v^{α_δ} несколько отличается от правила 2.3.1: если выполнено неравенство

$$I_\delta(v^{\alpha_2}) = f\left[\rho(A_h Bv^{\alpha_2}, u_\sigma)\right] \geqslant Cf\left[\Psi(\delta, \Omega(v^{\alpha_1}))\right] = C\bar{\Pi}_\delta(v^{\alpha_1}), \tag{3}$$

то в качестве приближенного решения задачи (1) выбирается такая экстремаль v^{α_δ} сглаживающего функционала, отвечающего найденному из уравнения (3.2) параметру регуляризации $\alpha_\delta > 0$, для которой справедливо неравенство

$$\rho(A_h Bv^{\alpha_\delta}, u_\sigma) \leqslant \Psi(\delta, \Omega(v^{\alpha_\delta})), \tag{4}$$

если же

$$I_\delta(v^{\alpha_2}) \leqslant C\bar{\Pi}_\delta(v^{\alpha_1}), \tag{5}$$

то v^{α_δ} выбирается так, чтобы выполнялось неравенство

$$\rho(A_h Bv^{\alpha_\delta}, u_\sigma) \geqslant \Psi(\delta, \Omega(v^{\alpha_\delta})). \tag{6}$$

Алгоритм *модифицированного обобщенного принципа квазирешений* (м.о.п.к.) для рассматриваемой задачи также аналогичен приведенному в § 3 с той лишь разницей, что в правиле отбора 2.3.2 вместо неравенств (3.3), (3.5) используются (3), (5).

Наконец, алгоритм *модифицированного обобщенного принципа сглаживающего функционала* (м.о.п.с.ф.) для задачи (1) отличается от приведенного в § 3 тем, что в правиле отбора используется неравенство

$$M^{\alpha_\delta}\left[v^{\alpha_\delta}\right] \geqslant f\left[C\Psi(\delta, \Omega(v^{\alpha_\delta}))\right].$$

Пусть теперь $\alpha_\delta > 0$ и $z^{\alpha_\delta} = Bv^{\alpha_\delta}$ получены по одному из модифицированных алгоритмов для решения разрешимого уравнения (1).

Теорема 1. *Справедливы оценки*

$$\|v^{\alpha_\delta}\| \leqslant R, \quad \rho(A_h z^{\alpha_\delta}, u_\sigma) \leqslant C\left[\sigma + \psi(h, \|B\| \cdot R)\right],$$

где $R = g^{-1}(\bar{\Omega}_1)$ *для алгоритмов м.о.п.н., м.о.п.к. и* $R = g^{-1}(\bar{\Omega}_2)$ *для алгоритма м.о.п.к.*

Доказательство этой теоремы проводится повторением рассуждений, приведенных в теоремах 1.13.1—1.13.6 с формальной заменой фигурирующих там величин $J_0^* = \mu_0$, λ_δ нулями, а также с заменой множителя 2, стоящего там перед величинами Ψ, на единицу.

Отметим, что, как и в § 1.13, для получения оценок используется лишь линейность и ограниченность оператора B. Таким образом, эти оценки верны и в случае, когда B — единичный оператор из Z в Z.

Для исследования вопроса об оптимальности порядка точности модифицированных алгоритмов введем множество корректности задачи (1) $M_R = \{z \in D\colon\ z = Bv,\ \|v\| \leqslant R\}$, где число $R > 0$ фиксировано и удовлетворяет условию $R \geqslant \|\bar{v}\|$. Рассмотрим множество разрешимых на D операторных уравнений вида

$$\bar{A}z = \bar{u}, \quad z \in D, \tag{7}$$

с непрерывными, взаимно однозначными операторами $\bar{A}$, действующими из D в U. Определим для данного R и фиксированных приближенных данных (A_h, u_σ) задачи (1) множество Σ_R точных данных $(\bar{A}, \bar{u})$ задач (7), имеющих одни и те же приближенные данные (A_h, u_σ) и имеющих решение $\bar{z} = \bar{z}(\bar{A}, \bar{u}) \in D$, которое подчинено условию $\bar{z} \in M_R$:

$$\Sigma_R \equiv \Sigma_R(A_h, u_\sigma, h, \sigma) =$$
$$= \{(\bar{A}, \bar{u})\colon\ \bar{A}\bar{z} = \bar{u},\quad \bar{z} \in M_R;\quad |\rho(\bar{A}z, \bar{u}) - \rho(A_h z, u_\sigma)| \leqslant CH \quad \forall z \in M_R\}.$$

Здесь $H \equiv \sigma + \psi(h, \|B\| \cdot R)$.

Определение 1. Будем называть *методом приближенного решения уравнения* (1) любой оператор P, ставящий в соответствие произвольному набору приближенных данных $(A_h, u_\sigma, h, \sigma)$ задачи (1) элемент $z_\delta \equiv P(A_h, u_\sigma, h, \sigma) \in D$.

Ясно, что метод P приближенного решения задачи (1) будет и методом решения задач (7) с точными данными из Σ_R. Поэтому введем характеристику точности метода P решения задачи (1) на приближенных данных $(A_h, u_\sigma, h, \sigma)$ следующим образом:

$$\Delta_P(A_h, u_\sigma, h, \sigma;\ R) = \sup_{\bar{A}, \bar{u}} \{\|P(A_h, u_\sigma, h, \sigma) - \bar{z}(\bar{A}, \bar{u})\|\colon\ (\bar{A}, \bar{u}) \in \Sigma_R\}. \tag{8}$$

Как и в § 1.13, дадим определения оптимальной точности и оптимального по порядку точности метода решения задачи (1).

Определение 2. *Оптимальной точностью* приближенного решения задачи (1) на данных $(A_h, u_\sigma, h, \sigma)$ назовем число

$$\Delta_{\text{опт}} = \Delta_{\text{опт}}(A_h, u_\sigma, h, \sigma;\ R) = \inf\{\Delta_P(A_h, u_\sigma, h, \sigma;\ R)\colon\quad P \in \mathcal{P}\},$$

где $\mathcal{P}$ — класс всех возможных методов решения задачи (1).

Определение 3. Метод P_0 решения задачи (1) имеет *оптимальный порядок точности* (*оптимален по порядку*) на данных $(A_h, u_\sigma, h, \sigma)$, если $\|P_0(A_h, u_\sigma, h, \sigma) - \bar{z}_0\| \leqslant k\,\Delta_{\text{опт}}(A_h, u_\sigma, h, \sigma;\ R)$ при $0 \leqslant h \leqslant h_0$, $0 \leqslant \sigma \leqslant \sigma_0$, причем число $k > 0$ не зависит от $(A_h, u_\sigma, h, \sigma)$ и от R.

Далее, как и в § 1.13, можно ввести оценочную функцию

$$\Theta(H, R) \equiv \Theta(H, R;\ A_h, u_\sigma, h, \sigma) =$$
$$= \sup_{\substack{\bar{A}_1, \bar{u}_1 \\ \bar{A}_2, \bar{u}_2}} \{\|\bar{z}(\bar{A}_1, \bar{u}_1) - \bar{z}(\bar{A}_2, \bar{u}_2)\|\colon (\bar{A}_{1,2}, \bar{u}_{1,2}) \in \Sigma_R\}. \tag{9}$$

Тогда будет справедлив аналог леммы 1.13.3.

Лемма 1. $\Delta_{\text{опт}} \geqslant \Theta(H, R)/2$.

Введем также аналог оценочной функции (1.13.28), имеющей в нашем случае вид

$$\omega_\delta(\tau, R) = \sup_{z_1, z_2} \{\| z_1 - z_2 \| : \rho(A_h z_{1,2}, u_\sigma) \leqslant \tau, \ z_{1,2} \in M_R\}, \tag{10}$$

которая будет определена при $\tau \geqslant H$. При этом можно уточнить результат леммы 1.13.4 для рассматриваемого случая разрешимых уравнений.

Лемма 2. $\Theta(H, R) \geqslant \omega_\delta(CH, R)$.

Действительно, как и при доказательстве леммы 1.13.4, найдутся такие элементы $z_{1,2} \in M_R$, что $\| z_1 - z_2 \| = \omega_\delta(CH, R)$, $\rho(A_h z_{1,2}, u_\sigma) \leqslant CH$. Рассмотрим пары (A_h, u_1), (A_h, u_2), где $u_{1,2} \equiv A_h z_{1,2}$. Поскольку по построению $\rho(u_{1,2}, u_\sigma) \leqslant CH$, то

$$z_{1,2} = \bar{z}(A_h, u_{1,2});$$

$$\| \rho(A_h z, u_{1,2}) - \rho(A_h z, u_\sigma) | \leqslant \rho(u_{1,2}, u_\sigma) \leqslant CH.$$

Таким образом, пары $(A_h, u_{1,2})$, рассматриваемые как точные данные задачи типа (7), принадлежат множеству Σ_R. Поэтому $\Theta(H, R) \geqslant \| z_1 - z_2 \| = \omega_\delta(CH, R)$.

Из лемм 1, 2 вытекает

Следствие 1. $\Delta_{\text{опт}} \geqslant \omega_\delta(CH, R)/2$.

Выберем теперь в качестве R величины, функционирующие в теореме 1. Тогда очевидно, что $\|\bar{v}\| \leqslant R$, $\rho(A_h \bar{z}, u_\sigma) \leqslant H \leqslant CH$, а по теореме 1

$$\| v^{\alpha_\delta} \| \leqslant R, \quad \rho(A_h z^{\alpha_\delta}, u_\sigma) \leqslant CH.$$

Тем самым $z^{\alpha_\delta} = Bv^{\alpha_\delta} \in M_R$, $\bar{z} = B\bar{v} \in M_R$ и по определению (10) функции и ω_δ с учетом следствия 1 справедлива оценка

$$\| z^{\alpha_\delta} - \bar{z} \| \leqslant \omega_\delta(CH, R) \leqslant 2\Delta_{\text{опт}}.$$

Итак, доказана

Теорема 2. *Модифицированные алгоритмы о.п.н., о.п.к., о.п.с.ф. для решения уравнения* (1) *в случае его разрешимости имеют оптимальный порядок точности на приближенных данных* $(A_h, u_\sigma, h, \sigma)$ *с константой оптимальности* $k = 2$.

Необходимо подчеркнуть также, что вследствие общей леммы 1.13.7 имеет место сходимость $\Delta_{\text{опт}}(A_h, u_\sigma, h, \sigma; R) \to 0$ при $\delta = (h, \sigma) \to 0$ для каждого $R > 0$.

Отметим также, что свойством оптимальности по порядку точности обладают и рассмотренные в § 5 алгоритмы о.м.н., м.к.

Замечание. Вопросы оптимальности по порядку точности модифицированных алгоритмов о.п.н., о.п.к. для линейных операторных уравнений с несколько иных позиций рассмотрены в [28, 88, 185, 187, 188]. Основным отличием подхода к исследованию оптимальности по порядку точности этих алгоритмов в общем нелинейном случае является использование других оценочных функций и, как следствие, другой техники доказательств.

§ 7. Сравнение априорного и апостериорного выбора параметра регуляризации

Предположим, что выполнены условия постановки задачи из § 1, и будем рассматривать задачу о приближенном нахождении Ω-оптимальных квазирешений уравнения (1.1) на множестве D. Для ее решения воспользуемся подходом, изложенным в § 1. При этом используются элементы $z^{\alpha} \in D$, реализующие точную нижнюю грань сглаживающего функционала на множестве D. Процедура построения приближений к множеству $\bar{Z}$ Ω-оптимальных квазирешений задачи (1.1) состоит из выбора параметра регуляризации $\alpha_\delta = \alpha(A_h, u_\sigma, h, \sigma)$ по приближенным данным задачи, из фиксации множества Z^{α_δ} экстремалей сглаживающего функционала для найденного α_δ и в выборе из Z^{α_δ} некоторого элемента z^{α_δ}, который и принимается в качестве приближения к множеству $\bar{Z}$.

Очевидно, весьма существенной частью этого алгоритма является выбор параметра регуляризации α_δ. Фактически на этом выборе построена идея вариационной регуляризации операторных уравнений с помощью сглаживающего функционала. С формальной точки зрения параметр регуляризации не обязательно должен выбираться по всей совокупности $(A_h, u_\sigma, h, \sigma)$ приближенных данных задачи. Возможен, например, вариант, когда величина α_δ выбирается только в зависимости от чисел (h, σ), определяющих погрешности приближенного оператора и правой части уравнения (1.1). Такой выбор параметра регуляризации $\alpha_\delta = \alpha(h, \sigma)$ называется *априорным*.

Иногда к априорному способу выбора параметра регуляризации относят также и вариант, в котором α_δ определяется лишь числами h, σ, $\|A_h\|$, $\|u_\sigma\|$, но не самими приближенными реализациями (A_h, u_σ) данных задачи.

Противоположная ситуация, когда α_δ выбирается не только по (h, σ), но и по величинам (A_h, u_σ), т. е. когда $\alpha_\delta = \alpha(A_h, u_\sigma, h, \sigma)$, называется *апостериорным выбором* параметра регуляризации.

Особенно заманчивой кажется возможность выбора параметра регуляризации только по приближенным реализациям (A_h, u_σ) без использования величин h, σ, оценивающих погрешности этих данных. Однако, как показано в [18, 20], даже для случая линейного, ограниченного, взаимно однозначного, точно заданного оператора A, действующего из бесконечномерного гильбертова пространства Z в гильбертово пространство U и не имеющего ограниченного обратного оператора, никакой выбор параметра регуляризации в виде $\alpha_\delta = \alpha(u_\sigma)$ $(\delta = (0, \sigma))$ не может дать регуляризующий алгоритм. Таким образом, использование совокупности величин $\delta = (h, \sigma)$ при построении регуляризующего алгоритма является принципиально необходимым требованием. Отметим, однако, что в случае решения линейных уравнений, с нормально разрешимым оператором и, в частности, систем линейных алгебраических уравнений возможно построение регуляризующего алгоритма при выборе параметра регуляризации лишь в зависимости от оценки h погрешности оператора: $\alpha = \alpha(h)$, без использования величины σ (см. [28, 55, 163, 164] и гл. 5).

Обратимся к рассмотрению регуляризующих алгоритмов для задачи (1.1), основанных на априорном выборе параметра регуляризации. Для задачи поиска Ω-оптимальных квазирешений на множестве D определим сглаживающий функционал в виде

$$M^\alpha[z]=\alpha\Omega(z)+K_\delta(z), \quad z\in D, \quad \alpha>0,$$

где $K_\delta(z)$ — модифицированный функционал невязки:

$$K_\delta(z)\equiv f[\rho(A_h z, u_\sigma)+\Psi(\delta, \Omega(z))],$$

а $\Psi(h, \Omega)=\sigma+\psi(h, \Omega)$ — полная мера аппроксимации уравнения (1.1). Вспомогательная функция $f(x)$, определенная при $x\geqslant 0$, предполагается обладающей следующими свойствами: $f(x)\in C^1\,[0, +\infty)$, монотонно возрастает и $f(+\infty)=+\infty$, $f(0)=0$.

Использование такого сглаживающего функционала допустимо в схеме решения экстремальных задач из § 1.4, так как функционал $\rho(A_h z, u_\sigma)+\Psi(\delta, \Omega(z))$ аппроксимирует невязку $\rho(Az, \bar{u})$ с мерой аппроксимации $2\Psi(\delta, \Omega)$ (ср. § 1.13). При этом справедлива оценка

$$K_\delta(z)\geqslant\inf\{K_\delta(z)\colon z\in D\}=\lambda_\delta\geqslant\mu_0, \tag{1}$$

которая будет существенно использоваться ниже. Введем экстремальную задачу: найти элементы $z^\alpha\in D$, для которых при данном $\alpha>0$ и фиксированном δ

$$M^\alpha[z^\alpha]=\inf\{M^\alpha[z]\colon z\in D\}. \tag{2}$$

В силу предположений 1—4 из § 1 и вследствие результатов из § 1.5 задача (2) имеет при каждом $\alpha>0$ непустое множество решений Z^α с элементами $z^\alpha\in D$.

Лемма 1. *Для любого $z^\alpha\in Z^\alpha$ справедливо неравенство*

$$\Omega(z^\alpha)\leqslant\bar{\Omega}+2f'[\mu_0+2\theta\Psi(\delta, \bar{\Omega})]\cdot[\Psi(\delta, \bar{\Omega})/\alpha], \quad 0<\theta<1.$$

Доказательство. Из экстремальности элемента z^α, условия аппроксимации (1.7), определения обобщенной меры несовместности λ_δ и из монотонности функции f следует, что

$$\begin{aligned}\alpha\Omega(z^\alpha)+f(\lambda_\delta)&\leqslant\alpha\Omega(z^\alpha)+f[\rho(A_h z^\alpha, u_\sigma)+\Psi(\delta, \Omega(z^\alpha))]=\\ &=M^\alpha[z^\alpha]\leqslant M^\alpha[\bar{z}]=\alpha\bar{\Omega}+f[\rho(A_h\bar{z}, u_\sigma)+\Psi(\delta, \bar{\Omega})]\leqslant\\ &\leqslant\alpha\bar{\Omega}+f[\rho(A\bar{z}, \bar{u})+2\Psi(\delta, \bar{\Omega})]=\alpha\bar{\Omega}+f[\mu_0+2\Psi(\delta, \bar{\Omega})].\end{aligned} \tag{3}$$

Учитывая при этом неравенство (1), получаем по формуле конечных приращений

$$\alpha[\Omega(z^\alpha)-\bar{\Omega}]\leqslant f[\mu_0+2\Psi(\delta, \bar{\Omega})]-f(\mu_0)=f'[\mu_0+2\theta\Psi(\delta, \bar{\Omega})]\cdot 2\Psi(\delta, \bar{\Omega}).$$

Это и доказывает лемму.

Теорема 1. *Пусть параметр регуляризации $\alpha_\delta=\alpha(h, \sigma)$ выбирается априорным образом так, чтобы при $\|\delta\|\to 0$ выполнялись условия $\alpha_\delta\to 0$, $\sigma/\alpha_\delta\to 0$, $\psi(h, \bar{\Omega})/\alpha_\delta\to 0$. Пусть также z^{α_δ} — произвольное решение задачи (2) при $\alpha=\alpha_\delta$. Тогда для любой последовательности $\|\delta_n\|\to 0$ при $n\to\infty$ соответствующая последовательность элементов*

$\{z_n\}$, *где* $z_n \equiv z^{\alpha_{\delta_n}}$, *обладает следующими свойствами:* $z_n \xrightarrow{\Sigma} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n \to \infty$.

Доказательство. Из леммы 1 вследствие условий выбора параметра α_δ и в силу сходимости $\Psi(\delta, \bar{\Omega}) = \sigma + \psi(h, \bar{\Omega}) \to 0$ при $\|\delta\| \to 0$ ясно, что

$$\overline{\lim_{\delta\to 0}}\, \Omega(z^{\alpha_\delta}) \leqslant \bar{\Omega} + 2\, \overline{\lim_{\delta\to 0}}\, f'\left[\mu_0 + 2\theta\Psi(\delta, \bar{\Omega})\right]\left[\sigma + \psi(h, \bar{\Omega})\right]/\alpha_\delta = \bar{\Omega}. \tag{4}$$

Из оценки (3) следует, что

$$f\left[\rho(A_h z^{\alpha_\delta}, u_\sigma)\right] \leqslant \alpha_\delta\left[\bar{\Omega} - \Omega^*\right] + f\left[\mu_0 + 2\Psi(\delta, \bar{\Omega})\right]. \tag{5}$$

Кроме того, согласно условию аппроксимации (1.7), вследствие вытекающей из доказанного соотношения (4) ограниченности величин $\Omega(z^{\alpha_\delta})$ справедлива оценка

$$\rho(A_h z^{\alpha_\delta}, u_\sigma) \geqslant \rho(A z^{\alpha_\delta}, \bar{u}) - \Psi(\delta, \Omega(z^{\alpha_\delta})) \geqslant \mu_0 - \Psi(\delta, \Omega_{\max}), \quad \Omega_{\max} \geqslant \Omega(z^{\alpha_\delta}). \tag{6}$$

Из (5), (6) с учетом монотонности и непрерывности функции f, а также сходимостей $\alpha_\delta \to 0$, $\Psi(\delta, \bar{\Omega}) \to 0$, $\Psi(\delta, \Omega_{\max}) \to 0$ при $\delta \to 0$ получим

$$\lim_{\delta\to 0} \rho(A_h z^{\alpha_\delta}, u_\sigma) = \mu_0. \tag{7}$$

Соотношения (4), (7) вместе с равенствами (1.6) позволяют использовать лемму 1.5.1, из которой и следуют доказываемые предельные свойства последовательности $\{z_n\}$.

Таким образом, априорный выбор параметра регуляризации, данный в теореме 1, обеспечивает сходимость приближений к множеству Ω-оптимальных квазирешений рассматриваемой задачи и не требует специального отбора экстремалей задачи (2). Для нахождения приближенного решения нет также необходимости вычислять обобщенную меру несовместности λ_δ.

Вместе с тем алгоритм, основанный на априорном выборе параметра регуляризации, в общем случае не гарантирует, что разность $\rho(A_h z^{\alpha_\delta}, u_\sigma) - \mu_0$ будет иметь тот же порядок малости (по σ, h), что и величины $\sigma + \psi(h, \Omega(z^{\alpha_\delta}))$ или $\sigma + \psi(h, \bar{\Omega})$. Это может привести к отсутствию оптимальности порядка точности для алгоритма с априорным выбором параметра регуляризации.

Такого недостатка заведомо лишены алгоритмы с апостериорным выбором параметра регуляризации, изученные в § 2, 3. Так, например, из результатов, полученных в § 1.13, следует справедливость оценки

$$\rho(A_h z^{\alpha_\delta}, u_\sigma) \leqslant \lambda_\delta + 2C\left[\sigma + \psi(h, \|B\| \cdot R)\right]$$

в случае поиска Ω-оптимального квазирешения уравнения (1.1). Для разрешимого операторного уравнения (1.1), согласно теореме 6.1, имеем

$$\rho(A_h z^{\alpha_\delta}, u_\sigma) \leqslant C\left[\sigma + \psi(h, \|B\| \cdot R)\right].$$

Число R в этих оценках определено, например, в теореме 6.1. Приведенные оценки обеспечивают оптимальность алгоритмов из § 2, 3 по порядку точности.

Если вспомогательные функции (2.1), определенные в § 2, 3 непрерывны при $\alpha>0$, то алгоритм о.п.н. гарантирует выполнение равенства

$$\rho(A_h z^{\alpha_\delta}, u_\sigma)=\lambda_\delta+\sigma+\psi(h, \Omega(z^{\alpha_\delta})),$$

а алгоритм м.о.п.н.— равенства

$$\rho(A_h z^{\alpha_\delta}, u_\sigma)=\sigma+\psi(h, \Omega(z^{\alpha_\delta})).$$

Такие равенства являются естественными условиями согласования невязки и погрешностей (h, σ) данных задачи, так как всякое квазирешение (решение) $\bar{z}$ уравнения (1.1) удовлетворяет неравенству

$$\rho(A_h\bar{z}, u_\sigma)\leqslant\mu_0+\sigma+\psi(h, \Omega(\bar{z}))$$
$$(\rho(A_h\bar{z}, u_\sigma)\leqslant\sigma+\psi(h, \Omega(\bar{z}))).$$

Аналогичным образом априорный выбор параметра регуляризации не обеспечивает в общем случае выполнения равенства $\Omega(z^{\alpha_\delta})=\omega_\delta$ где ω_δ — известная оценка для $\bar{\Omega}$: $\omega_\delta\geqslant\bar{\Omega}$, $\omega_\delta\to\bar{\Omega}$ при $\|\delta\|\to 0$. В рассмотренном выше случае непрерывности вспомогательных функции (2.1) алгоритм о.п.к. дает равенство $\Omega(z^{\alpha_\delta})=\omega_\delta$.

Итак, алгоритмы с апостериорным выбором α не только обеспечивают сходимость приближений и сходимости по функционалам

$$\lim_{\delta\to 0}\Omega(z^{\alpha_\delta})=\bar{\Omega}, \quad \lim_{\delta\to 0}\rho(A_h z^{\alpha_\delta}, u_\sigma)=\mu_0,$$

но и дают необходимые значения некоторых функционалов (невязки стабилизирующего функционала и др.) на приближенном решении.

В заключение приведем краткую историческую справку об исследованиях по выбору параметра регуляризации для вариационных методов решения операторных уравнений.

Исторически первые работы в этой области были посвящены априорному выбору параметра регуляризации. В [190, 191] такой выбор был предложен для линейных совместных операторных уравнений, а в [192, 193]—для нелинейных уравнений. Априорный выбор параметра регуляризации для решения задач оптимального управления, оптимального планирования и задач оптимизации функционалов дан в [195, 196, 198, 199], а в [197] такой выбор параметра регуляризации предложен для решения систем линейных алгебраических уравнений.

В дальнейшем эта ветвь исследований по выбору параметра регуляризации была связана с изучением линейных и нелинейных операторных уравнений в различных функциональных и топологических пространствах, а также конкретных видов уравнений [11, 14, 25, 27—29, 45, 55, 88, 108, 142, 154, 160, 164, 181, 182, 187, 202]. Для решения экстремальных задач вопрос изучался в [30, 32, 33, 51, 142, 153, 203].

Первые работы по апостериорному выбору параметра регуляризации относятся к 1966 г. В [86] были предложены идеи выбора параметра регуляризации по *принципам невязки* и *квазирешений* для решения линейных задач с точно заданным оператором. Такой выбор α_σ сводился к решению уравнений

$$\beta(\alpha) \equiv \| Az^\alpha - u_\sigma \| = \sigma, \quad \gamma(\alpha) \equiv \| z^\alpha \| = \omega_\sigma.$$

Независимо принцип невязки был предложен и обоснован в работе [149], а затем в [151, 152]. С различных сторон принцип невязки изучался рядом авторов [16, 56, 63, 65, 82, 88, 139, 142, 159, 160, 163, 164, 187, 202]. Выбор параметра регуляризации по *принципу сглаживающего функционала* для решения линейных задач с точно заданным оператором был введен в работе [150] и затем развит в [151] для случая нелинейных задач с точно заданным оператором. Он сводился к нахождению параметра α_δ из уравнения

$$\varphi(\alpha) \equiv M^\alpha [z^\alpha] = C\sigma^2, \quad C = \text{const} > 1.$$

В дальнейшем такой выбор параметра регуляризации изучался в [141, 142].

Упомянутые апостериорные способы выбора α непосредственно не переносятся на общий нелинейный случай (см. § 1). Примеры и причины отсутствия непосредственного обобщения на этот случай принципа невязки и принципа квазирешений приведены в § 2 (см. также [63, 65]). Принцип невязки и сглаживающего функционала не применим даже в случае линейных уравнений с приближенно заданным оператором, если не использовать дополнительной априорной информации о норме искомого точного решения. Попытка обобщить принцип невязки и принцип сглаживающего функционала путем использования такой дополнительной априорной информации была предпринята в работах [142, 152] (см. также [160, 162, 164]).

Обобщение принципа невязки на случай приближенного совместного линейного операторного уравнения без использования информации о норме точного решения было дано в [59, 61]. Именно там был введен *обобщенный принцип невязки*, в котором параметр α_δ выбирался из уравнений

$$\rho(\alpha) \equiv \| A_h z^\alpha - u_\sigma \|^2 - (\sigma + h \| z^\alpha \|)^2 - \mu_\delta^2 = 0, \tag{8}$$

где

$$\mu_\delta = \inf \{ \| A_h z - u_\sigma \| : z \in Z \}.$$

Применение обобщенного принципа невязки позволило впервые аккуратно учесть погрешность конечномерной аппроксимации линейных некорректных задач при их численном решении [62]. При этом не использовалась информация типа оценок нормы точного решения.

В дальнейшем изучение обобщенного принципа невязки было предпринято в работах [88, 186, 187], где было показано, что в случае разрешимых линейных операторных уравнений число μ_δ в формуле (8) может быть опущено. Рассмотрен был также случай, когда о.п.н. формально дает $\alpha_\delta = 0$.

Обобщенный принцип невязки изучался в различных функциональных пространствах весьма общего вида [104, 105, 206, 222, 223, 225, 226]. Рассматривался также его итеративный вариант [19, 28]. Обобщенный принцип невязки использовался в линейном случае для поиска L-псевдорешений [3, 54, 55, 164].

В связи с разрывностью функций $\rho(\alpha)$ обобщенный принцип невязки непосредственно не может быть использован в общем случае для решения нелинейных операторных уравнений. Попытки приспособить его и в этом случае делались в работах [161, 162], где делались дополнительные предположения о существовании обычного решения уравнения (8) (принцип оптимальности невязки). Предложен был также принцип сравнения невязки, позволявший осуществлять некоторый специальный отбор приближенного решения z^{α_δ} (см. [13]) при поиске нормальных решений нелинейных уравнений.

Обобщенный принцип невязки для решения нелинейных приближенно заданных операторных уравнений (в том числе и несовместных) предложен в работах [119, 120, 123—125, 130]. В работе [130] введены также *обобщенные принципы квазирешений* и *сглаживающего функционала*, позволяющие выбирать параметр регуляризации и приближенное решение в общем нелинейном случае.

В частности, алгоритм о.п.с.ф. не использует информацию о норме искомого решения в отличии, например, от принципа сглаживающего функционала [142].

Отметим в связи с работами по апостериорному выбору параметра регуляризации также альтернативный принцип выбора α (см. [63, 65, 206]), применимый для совместных нелинейных операторных уравнений с точно заданным оператором, а также способ отбора приближенного решения нелинейных несовместных уравнений, данный в работе [29].

Иные подходы к проблеме выбора параметра регуляризации изучались в [6, 12, 18, 96, 118, 125, 204, 205] и в некоторых других работах.

§ 8. О задачах с избыточной информацией

Рассмотрим операторное уравнение вида

$$A[x, z] = \bar{u}(x), \tag{1}$$

в котором по правой части $\bar{u}(x) \in U = L_1[c, d]$ ищется неизвестное $z \in D$. Поставим для уравнения (1) задачу поиска Ω-оптимальных квазирешений на D, как это делалось в § 1. Данными для их приближенного нахождения будут служить величины (A_h, h, ψ), определенные в § 1, а также несколько функций, которые представляют собой приближения к $\bar{u}(x)$: $u_1(x), \ldots, u_n(x) \in L_1[c, d]$. В отличие, однако, от предположений § 1 эти функции, вообще говоря, не удовлетворяют введенному там условию аппроксимации $\| u_\sigma - \bar{u} \|_{L_1} \leqslant \sigma$.

Будем предполагать, что для каждого $x \in [c, d]$ (кроме, возможно, точек из некоторого множества E меры нуль) не менее половины функций $u_k(x)$ $(k = 1, \ldots, n)$ подчиняются неравенству $|u_k(x) - \bar{u}(x)| \leqslant \sigma$, а остальные функции из этого набора такому неравенству в точке x могут не удовлетворять. Требуется лишь конечность значений $u_k(x)$, $\bar{u}(x)$ при этом x. Число $\sigma \geqslant 0$ считается известным.

На практике подобная ситуация встречается, когда правая часть уравнения (1) получается в результате многократных измерений с возможными грубыми ошибками в некоторых точках x.

В данном параграфе будет указан способ, позволяющий находить по данным $\{u_k(x)\}_{k=1}^n$ некоторую функцию $\tilde{u}_\sigma(x)$, конечную всюду на $[c, d]\setminus E$, суммируемую и удовлетворяющую условию аппроксимации

$$|\tilde{u}_\sigma(x)-\bar{u}(x)|\leqslant\sigma \quad \forall x\in[c, d]\setminus E. \tag{2}$$

Таким образом, функция $\tilde{u}_\sigma(x)$ лишена недостатков каждой конкретной функции $u_k(x)$, выражающихся в нарушении условия (2) в ряде точек x. И если для функции u_k стандартное условие аппроксимации $\|u_k-\bar{u}\|\leqslant(d-c)\sigma$, вообще говоря, не выполнено, то для $\tilde{u}_\sigma(x)$ оказывается справедливой оценка

$$\|\tilde{u}_\sigma(x)-\bar{u}(x)\|_{L_1[c, d]}\leqslant(d-c)\sigma.$$

Поэтому найденная функция $\tilde{u}_\sigma(x)$ может быть использована для нахождения Ω-оптимальных квазирешений уравнения (1) по алгоритмам из § 2—7.

Перейдем к изложению результатов. Определим на $L_1[c, d]$ функционал

$$\Phi(u)=\sum_{k=1}^n\|u-u_k\|_{L_1}=\int_c^d\left\{\sum_{k=1}^n|u(x)-u_k(x)|\right\}dx$$

и изучим задачу минимизации: найти такую функцию $\tilde{u}(x)\in L_1[c, d]$, для которой

$$\Phi(\tilde{u})=\inf\{\Phi(u)\colon u\in L_1[c, d]\}. \tag{3}$$

Для этой задачи справедлива

Теорема 1. *Задача* (3) *разрешима. Любое ее решение* $\tilde{u}(x)$ *конечно всюду, кроме, быть может, множества* E, *и удовлетворяет неравенству*

$$|\tilde{u}(x)-\bar{u}(x)|\leqslant\sigma \quad \forall x\in[c, d]\setminus E.$$

Для доказательства понадобится следующая лемма.

Лемма 1. *Если задан набор чисел* $\{y_k\}_{k=1}^n$ *такой, что не менее половины из них удовлетворяет неравенству* $|y_k-\bar{y}|\leqslant\sigma$, *то любое решение задачи: найти такое число* $\tilde{y}$, *для которого*

$$\sum_{k=1}^n|y_k-\tilde{y}|=\inf\left\{\sum_{k=1}^n|y-y_k|\colon y\in\mathbf{R}\right\}, \tag{4}$$

лежит на отрезке $[\bar{y}-\sigma,\ \bar{y}+\sigma]$.

Доказательство. Введем функцию

$$f_0(y)=\sum_{k=1}^n|y-y_k|.$$

Эта функция непрерывна и выпукла, причем $f_0(y)\to+\infty$ при $y\to\infty$. Поэтому задача (4) разрешима. Покажем, что ее решения могут лежать лишь на отрезке $[\bar{y}-\sigma, \bar{y}+\sigma]$. Для этого заметим, что функция $f_0(y)$ имеет всюду, кроме точек y_k, производную $f'_0(y)$, которая вычисляется по правилу $f'_0(y)=-n+2\nu(y)$. Здесь $\nu(y)$—число точек из набора $\{y_k\}$, для которых выполнено неравенство $y>y_k$. Производная $f'_0(y)$ монотонно не убывает. Известно [26], что необходимым условием минимума функции $f_0(y)$ с указанными свойствами, достигающегося в точке $\tilde{y}$, является требование $f'_0(\tilde{y})=0$. Последнее уравнение понимается как уравнение с монотонной функцией (см. § 1.6).

Вне отрезка $[\bar{y}-\sigma, \bar{y}+\sigma]$ содержится менее половины точек из набора $\{y_k\}$. Поэтому при $y<\bar{y}-\sigma$ справедливо неравенство $\nu(y)<n/2$, так что $f'_0(y)<0$. Аналогичным образом в области $y<\bar{y}+\sigma$ лежит более половины точек набора, так что $\nu(y)>n/2$ при $y>\bar{y}+\sigma$ и $f'_0(y)>0$. Из сказанного ясно, что решения уравнения с монотонной функцией $f'_0(y)=0$ могут лежать лишь на отрезке $[\bar{y}-\sigma, \bar{y}+\sigma]$. Лемма доказана.

Отметим, что процесс решения задачи (4), основанный на решении уравнения $f'_0(y)=0$, легко реализуется численно.

Доказательство теоремы 1. Рассмотрим задачу типа (4) при фиксированном $x\notin E$ найти число $\tilde{u}(x)$ такое, что

$$\sum_{k=1}^{n}|\tilde{u}(x)-u_k(x)|=\inf\left\{\sum_{k=1}^{n}|u(x)-u_k(x):\ u(x)\in\mathbf{R}\right\}. \tag{5}$$

По лемме 1 решение этой задачи существует. Тогда при $x\in[c, d]\setminus E$ задана некоторая функция $\tilde{u}(x)$, которая в силу (5) удовлетворяет неравенству

$$\sum_{k=1}^{n}|\tilde{u}(x)-u_k(x)|\leqslant\sum_{k=1}^{n}|\bar{u}(x)-u_k(x)|\equiv\varphi(x). \tag{6}$$

По условиям функция $\varphi(x)\in L_1[c, d]$. Согласно известному свойству интегрируемых по Лебегу функций (см. теорему 0.4.7), из оценки (6) следует интегрируемость по Лебегу и функции, стоящей в оценке (6) слева. Тогда из (5) получим

$$\Phi(\tilde{u})=\int_a^b\left\{\sum_{k=1}^{n}|\tilde{u}(x)-u_k(x)|\right\}dx\leqslant\Phi(u)=\int_a^b\left\{\sum_{k=1}^{n}|u(x)-u_k(x)|\right\}dx,$$

причем это неравенство справедливо для любой функции $u(x)\in$ $\in L_1[c, d]$.

Таким образом, функция $\tilde{u}(x)$ есть решение задачи (3). По лемме 1 и в силу (5) будет выполнена оценка $|\tilde{u}(x)-\bar{u}(x)|\leqslant\sigma$ для любого $x\in[c, d]\setminus E$. Теорема доказана.

С алгоритмической точки зрения нахождение функции $\tilde{u}_\sigma(x)\equiv\tilde{u}(x)$ связано с многократным решением задачи типа (4).

Другие подходы к задачам с избыточной информацией изучались в [108] и др.

ГЛАВА 3

КОНЕЧНОМЕРНЫЕ ВАРИАНТЫ АЛГОРИТМОВ

§ 1. Конечномерная аппроксимация нелинейных некорректных задач

Пусть (Z, τ) — хаусдорфово топологическое пространство, на некотором подмножестве $D (D \neq \varnothing)$ которого задан функционал $J(z)$. Введем задачу минимизации функционала $J(z)$ по аргументу на множестве D: найти элементы $z^* \in D$, для которых

$$J(z^*) = \inf\{J(z) \colon z \in D\} \equiv J^*. \tag{1}$$

Предположим, что Z^* — непустое множество решений задачи (1), и рассмотрим задачу отыскания Ω-оптимальных решений: найти элементы $\bar{z} \in Z^*$ такие, что

$$\Omega(\bar{z}) = \inf\{\Omega(z^*) \colon z^* \in Z^*\} \equiv \bar{\Omega}.$$

При этом мы считаем, что функционал $\Omega(z)$ задан на множестве D и ограничен там снизу: $\inf\{\Omega(z) \colon z \in D\} \equiv \Omega^* > -\infty$. Будем предполагать, что точный функционал задачи (1) неизвестен, а исходными данными для приближенного нахождения Ω-оптимальных решений являются величины $\{J_\delta, \delta, \Psi_0\}$. Здесь $J_\delta(z)$ — приближенный функционал задачи (1), принадлежащий множеству $\mathscr{F}$ допустимых приближенных функционалов, определенных на D, ограниченных там снизу и удовлетворяющих условию аппроксимации

$$|J(z) - J_\delta(z)| \leqslant \Psi_0(\delta, \Omega(z)) \quad \forall z \in D. \tag{2}$$

При этом известны мера аппроксимации Ψ_0 и числовой вектор $\delta \in \mathbf{R}^k_+$ $(\delta \neq \theta = (0, \ldots, 0) \in \mathbf{R}^k)$.

Рассматриваемая таким образом о с н о в н а я з а д а ч а приближенного нахождения Ω-оптимальных решений заключается в построении элемента $z_\delta \equiv z(J_\delta, \delta, \Psi_0) \in D$ такого, что z_δ τ-секвенциально сходится к множеству $\bar{Z}$ Ω-оптимальных решений задачи (1) при $\|\delta\| \to 0$. Эта задача рассматривалась в гл. 1. Будем считать, что функционалы J, J_δ, Ω и функция Ψ_0 удовлетворяют требованиям, изложенным в § 1.2.

Решение основной задачи, как правило, невозможно без ЭВМ Для ее решения с помощью ЭВМ, вообще говоря, необходимо произвести конечномерную аппроксимацию задачи. Во многих случаях в качестве исходных данных имеются именно конечномерные величины, в некотором смысле аппроксимирующие функционалы J_δ. В данной главе будет изучаться задача построения конечномерных приближений решения основной задачи. Для этих целей будут использоваться конечномерные аналоги алгоритмов обобщенных принципов невязки, квазирешений и сглаживающего функционала [135].

Остановимся на постановке задачи конечномерной аппроксимации. Пусть Z_N — конечномерное нормированное пространство размерности N с элементами, которые будут обозначаться символом $\hat{z}_N$ (например, $\hat{z}^*_N \in Z_N$). Введем операторы P_N, $\bar{P}_N$, осуществляющие отображения P_N: $Z \to Z_N$, $\bar{P}_N$: $Z_N \to Z$. Операторы P_N, $\bar{P}_N$ должны обладать рядом свойств, среди которых перечислим пока следующие:

1) операторы P_N, $\bar{P}_N$ непрерывны при любом N;

2) $P_N \bar{P}_N = E_N$, где E_N — единичный оператор в пространстве Z_N;

3) для любого $z \in D$ и каждого N справедливо включение $\bar{P}_N P_N z \in D$

Введем множества $\hat{D}_N \equiv \hat{D} = \{\hat{z}_N\colon \bar{P}_N \hat{z}_N \in D\}$. В силу свойства 3) операторов P_N, $\bar{P}_N$ множества $\hat{D}_N \subset Z_N$ непусты, так как содержат элементы вида $\hat{z}_N = P_N z$ $(z \in D)$. Множества $\hat{D}_N$ будут представлять собой конечномерные аналоги множества D. В дальнейшем для краткости будем опускать индекс N, определяющий размерность соответствующего конечномерного пространства Z_N, там, где этот индекс считается фиксированным (например, $\bar{P}\hat{z} \equiv \bar{P}_N \hat{z}_N$ и т. д.).

Предположим, что вместо функционала $J_\delta(z)$ в нашем распоряжении имеется его конечномерная аппроксимация, т. е. заданный на множестве $\hat{D}_N$ конечномерный функционал $\hat{J}_{\delta N}(\hat{z}) \equiv \hat{J}_\eta(\hat{z})$, удовлетворяющий следующему условию аппроксимации:

$$|\hat{J}_{\delta N}(P_N z) - J_\delta(z)| \leqslant \hat{\Psi}(\eta, \Omega(z)) \quad \forall z \in D. \tag{3}$$

Здесь вектор $\eta \in \mathbf{R}^{k+1}_+$ имеет вид $\eta = (\delta, 1/N)$, а функция $\hat{\Psi}$ есть известная мера аппроксимации функционала J_δ конечномерным функционалом $\hat{J}_\eta \equiv \hat{J}_{\delta N}$. Будем считать, что функция $\hat{\Psi}$ обладает теми же свойствами, что и функция Ψ_0.

Основная конечномерная задача. По набору данных $\{\hat{J}_\eta, \hat{\Psi}, \Psi_0, \delta, N, Z_N\}$ построить такой элемент $\hat{z}_\eta \in \hat{D}_N$, что семейство $\{z_\eta\}$, где $z_\eta \equiv \bar{P}_N \hat{z}_\eta \in D$, τ-секвенциально сходится к $\bar{Z}$ при $\eta \to 0$. Найденные таким образом элементы $z_\eta = \bar{P}_N \hat{z}_\eta$ принимаются в качестве решений основной задачи.

Изложим принятую в данной главе схему решения основной конечномерной задачи. Для этого определим на множестве $\hat{D}$ функционалы

$$\hat{\Omega}(\hat{z}) \equiv \Omega(\bar{P}\hat{z}), \quad \hat{I}(\hat{z}) \equiv \hat{I}_\eta(\hat{z}) = f[\hat{J}_{\delta N}(\hat{z})],$$

где $f(x) \in \mathscr{F}^m$ — фиксированная вспомогательная функция. Определим также конечномерный сглаживающий функционал

$$\hat{M}^\alpha[\hat{z}] = \alpha\hat{\Omega}(\hat{z}) + \hat{I}(\hat{z}), \quad \alpha > 0, \quad \hat{z} \in \hat{D}, \tag{4}$$

и введем для него экстремальную задачу: найти элементы $\hat{z}^\alpha \in \hat{D}$, для которых

$$\hat{M}^\alpha[\hat{z}^\alpha] = \inf\{\hat{M}^\alpha[\hat{z}]: \hat{z} \in \hat{D}\}. \tag{5}$$

Пусть $\hat{Z}^\alpha \neq \varnothing$ — множество ее решений. Для решения основной конечномерной задачи примем схему из § 1.4:

а) выбор параметра регуляризации $\alpha_\eta = \alpha\{\hat{J}_\eta, \hat{\Psi}, \Psi_0, \delta, N, Z_N\}$;

б) отбор решения $\hat{z}_\eta \equiv \hat{z}^{\alpha_\eta}$ задачи (5) при $\alpha = \alpha_\eta$ из множества $\hat{Z}^{\alpha_\eta}$ по некоторому правилу.

Методы выбора параметра регуляризации и отбора элемента $\hat{z}_\eta$, исследуемые в данной главе, будут представлять собой конечномерные аналоги алгоритмов, изученных в § 1.7—1.9. При этом кроме сформулированных уже в данном параграфе требований будут предполагаться выполненными следующие условия:

4) для любого $\bar{z} \in \bar{Z}$ $\overline{\lim\limits_{N \to \infty}}\, \Omega(\bar{P}_N P_N \bar{z}) \leqslant \bar{\Omega}$;

5) при каждом η функционал $\hat{J}_{\delta N}$ полунепрерывен снизу на множестве $\hat{D}$.

Перечисленные условия обеспечивают возможность реализации изложенной схемы решения основной конечномерной задачи. Для обоснования этого рассмотрим следующую лемму.

Лемма 1. *При каждом фиксированном N всякое непустое множество $\hat{\Omega}_C \equiv \{\hat{z} \in \hat{D}_N: \hat{\Omega}(\hat{z}) \leqslant C\}$ замкнуто и относительно компактно в Z_N.*

Доказательство. Рассмотрим произвольную последовательность $\{\hat{z}_n\} \subset \hat{\Omega}_C$ и соответствующую последовательность элементов $\{z_n\}$, где $z_n = \bar{P}\hat{z}_n$. Из определения множества $\hat{D}$ вытекает включение $\{z_n\} \subset D$, причем $\Omega(z_n) = \Omega(\bar{P}\hat{z}_n) = \hat{\Omega}(\hat{z}_n) \leqslant C$. Таким образом, $\{z_n\} \subset \Omega_C = \{z \in D: \Omega(z) \leqslant C\}$. Согласно предположению 2 из § 1.2, множество Ω_C будет τ-секвенциально компактным. Поэтому из $\{z_n\}$ можно извлечь τ-сходящуюся при $k \to \infty$ подпоследовательность $\{z_{n_k}\} \subset D$: $z_{n_k} \xrightarrow{\tau} z_0 \in \Omega_C$. Отсюда в силу непрерывности оператора P и по свойству 2) операторов P, $\bar{P}$ получим $Pz_{n_k} = P(\bar{P}\hat{z}_{n_k}) = \hat{z}_{n_k} \xrightarrow{Z_N} Pz_0 \equiv \hat{z}_0$ при $k \to \infty$. Эта сходимость вместе с непрерывностью оператора $\bar{P}$ дает $z_{n_k} = \bar{P}\hat{z}_{n_k} \xrightarrow{\tau} \bar{P}\hat{z}_0$. Поскольку пространство (Z, τ) хаусдорфово, то предел последовательности должен быть единственным (см. § 0.1). Поэтому $\bar{P}\hat{z}_0 = z_0 \in D$. Тем самым $\hat{z}_0 \in \hat{D}$, и вследствие включения $z_0 \in \Omega_C$ справедливо соотношение $\hat{\Omega}(\hat{z}_0) = \Omega(\bar{P}\hat{z}_0) = \Omega(z_0) \leqslant C$. Это означает, что $\hat{z}_0 \in \hat{\Omega}_C$.

Лемма 2. *При фиксированном η функционалы $\hat{\Omega}(\hat{z})$, $\hat{I}(\hat{z})$, $\hat{M}^\alpha[\hat{z}]$ полунепрерывны снизу на $\hat{D}$.*

Действительно, возьмем произвольную последовательность $\{\hat{z}_n\} \subset \hat{D}$ такую, что $\hat{z}_n \to \hat{z}_0 \in \hat{D}$ при $n \to \infty$. Тогда в силу непрерывности оператора $\bar{P}$ и по определению множества $\hat{D}$ имеем $\bar{P}\hat{z}_n \xrightarrow{\tau} \bar{P}\hat{z}_0 \in D$ при

$n\to\infty$, а вследствие τ-секвенциальной полунепрерывности снизу на множестве D функционала Ω (см. § 2.1) справедливо соотношение

$$\varliminf_{n\to\infty}\hat{\Omega}(\hat{z}_n)=\varliminf_{n\to\infty}\Omega(\bar{P}\hat{z}_n)\geqslant\Omega(\bar{P}\hat{z}_0)=\hat{\Omega}(\hat{z}_0).$$

Далее, из полунепрерывности снизу на $\hat{D}$ функционала $\hat{J}_{\delta N}$ и непрерывности и возрастания функции f получим

$$\varliminf_{n\to\infty}\hat{I}(\hat{z}_n)=\varliminf_{n\to\infty}f[\hat{J}_{\delta N}(\hat{z}_n)]\geqslant f[\hat{J}_{\delta N}(\hat{z}_0)]=\hat{I}(\hat{z}_0).$$

Эти соотношения вместе с равенством (4) доказывают лемму.

Леммы 1, 2 обосновывают выполнение для функционалов $\hat{\Omega}(\hat{z})$, $\hat{I}(\hat{z})$ предположений 1, 2 из § 1.2 на множестве $\hat{D}$ при фиксированных N, δ. При этом в качестве топологии в Z_N принимается топология сходимости по норме.

Из лемм 1, 2 вытекает, таким образом, ряд утверждений, аналогичных полученным в § 1.5, 1.6 с использованием предположений 1, 2 из § 1.2. В частности, верна

Теорема 1. *При каждом фиксированном* η *и для любого* $\alpha>0$ *экстремальная задача* (5) *разрешима, так что* $\hat{Z}^\alpha\neq\varnothing$.

Справедливость этой теоремы вытекает из теоремы 1.4.2, примененной к функционалу $\hat{M}^\alpha$ с учетом результатов лемм 1, 2.

Обращаясь еще раз к задаче (5), сделаем следующее замечание — эта задача конечномерная. Для решения подобных задач существует ряд эффективных методов [31, 91, 170—173, 215, 219]. Особенно удобен случай, когда множество $\hat{D}$ выпуклое и выпуклыми же на $\hat{D}$ являются функционалы $\hat{\Omega}$, $\hat{I}$. Тогда задача (5) становится задачей выпуклого программирования. Если $\hat{D}=Z_N$, а $\hat{\Omega}$, $\hat{I}$ — квадратичные функционалы, то задача (5) сводится к решению системы линейных уравнений [31, 91, 173].

В заключение отметим некоторые следствия условий аппроксимации (2), (3). Прежде всего для любого $z\in D$ выполнено неравенство

$$|J(z)-\hat{J}_\eta(Pz)|\leqslant\Psi_0(\delta,\Omega(z))+\hat{\Psi}(\eta,\Omega(z))\equiv\Psi(\eta,\Omega(z)). \tag{6}$$

Согласно свойству 3) операторов P, $\bar{P}$, для любого $\bar{z}\in\bar{Z}\subset D$ имеем $\bar{P}(P\bar{z})\in D$, так что элемент $\hat{\bar{z}}\equiv P\bar{z}$ принадлежит множеству $\hat{D}$. Поэтому из (6) следует

$$|J^*-\hat{J}_\eta(\hat{\bar{z}})|=|J(\bar{z})-\hat{J}_\eta(P\bar{z})|\leqslant\Psi(\eta,\Omega(\bar{z}))=\Psi(\eta,\bar{\Omega}). \tag{7}$$

Наконец, принимая во внимание свойство 2) операторов P, $\bar{P}$, получаем из (6) для каждого $\hat{z}\in\hat{D}$

$$|J(\bar{P}\hat{z})-\hat{J}_\eta(P\bar{P}\hat{z})|=|J(\bar{P}\hat{z})-\hat{J}_\eta(\hat{z})|\leqslant\Psi(\eta,\hat{\Omega}(\hat{z})). \tag{8}$$

Функция $\Psi(\eta,\Omega(z))$, введенная в (6), представляет собой совокупную меру аппроксимации точного функционала J конечномерным функционалом $\hat{J}_\eta$. Она, очевидно, обладает теми же свойствами, что и функция Ψ_0 (см. § 2.1).

Замечания. Конечномерная аппроксимация некорректных задач исследовалась многими авторами. В этой связи упомянем работы [2, 23, 24, 35—38, 51, 53, 88, 142, 160, 164, 183, 184, 187, 202, 204—206]. При этом не всегда производился учет погрешности конечномерной аппроксимации при выборе регуляризованного приближения, особенно при апостериорном выборе параметра регуляризации. Первая работа, в которой дан аккуратный учет погрешности конечномерной аппроксимации с помощью обобщенного принципа невязки, касалась решения линейных операторных уравнений [62]. В дальнейшей разработке этой тематики посвящались работы [80, 116, 117, 134, 135]. В частности, в [116] рассматривался случай нелинейных некорректных задач и использовался альтернативный принцип выбора параметра регуляризации. В работах [134, 135] рассматривалась конечномерная аппроксимация нелинейных задач на основе использования алгоритмов о.п.н., о.п.к., о.п.с.ф.

§ 2. Конечномерный обобщенный принцип невязки

Определим при фиксированном $\eta=(\delta, 1/N)$ величины

$$\hat{\lambda}\equiv\hat{\lambda}_\eta=\inf\{\hat{J}_\eta(\hat{z})+\Psi(\eta, \hat{\Omega}(\hat{z})):\ \hat{z}\in\hat{D}\},$$
$$\hat{\mu}\equiv\hat{\mu}_\eta=\inf\{\hat{J}_\eta(\hat{z}):\ \hat{z}\in\hat{D}\},$$
$$\hat{\Omega}^*\equiv\hat{\Omega}^*_N=\inf\{\hat{\Omega}(\hat{z}):\ \hat{z}\in\hat{D}\}=\inf\{\Omega(\bar{P}\hat{z}):\ \hat{z}\in\hat{D}\}.$$

Из этих определений легко получаются неравенства $\hat{\lambda}\geqslant\hat{\mu}$, $\hat{\Omega}^*\geqslant\Omega^*$. Кроме того, числа $\hat{\lambda}$ обладают следующим свойством.

Лемма 1. *При каждом фиксированном* η *справедливо неравенство* $\hat{\lambda}_\eta\geqslant J^*$, *а при* $\eta\to0$ *имеет место сходимость* $\hat{\lambda}_\eta\to J^*$.

Доказательство. Из неравенства (1.6) с использованием свойства 2) операторов P, $\bar{P}$ получим

$$J^*=\inf\{J(z):\ z\in D\}\leqslant J(\bar{P}\hat{z})\leqslant\hat{J}_\eta(P\bar{P}\hat{z})+\Psi(\eta, \Omega(\bar{P}\hat{z}))=$$
$$=\hat{J}_\eta(\hat{z})+\Psi(\eta, \hat{\Omega}(\hat{z}))\quad\forall\hat{z}\in\hat{D}. \qquad (1)$$

Здесь использована также вытекающая из определения множества $\hat{D}$ принадлежность $\bar{P}\hat{z}\in D$. Из (1) с учетом (1.7) следует

$$J^*\leqslant\inf\{\hat{J}_\eta(\hat{z})+\Psi(\eta, \hat{\Omega}(\hat{z})):\ \hat{z}\in\hat{D}\}=\hat{\lambda}_\eta\leqslant\hat{J}_\eta(\hat{\bar{z}})+\Psi(\eta, \hat{\Omega}(\hat{\bar{z}}))\leqslant$$
$$\leqslant J^*+\Psi(\eta, \bar{\Omega})+\Psi(\eta, \Omega(\bar{P}_N P_N\bar{z})). \qquad (2)$$

Используем далее свойство 4) операторов P, $\bar{P}$. Из него ясно, что при $N\geqslant N_0=\mathrm{const}$ выполнена оценка $\Omega(\bar{P}_N P_N\bar{z})\leqslant\tilde{\Omega}=\mathrm{const}$. Тогда из свойств функции Ψ получается соотношение

$$0\leqslant\Psi(\eta, \Omega(\bar{P}_N P_N\bar{z}))\leqslant\Psi(\eta, \tilde{\Omega})\to0,\quad \eta\to0.$$

Таким образом, имеют место сходимости $\Psi(\eta, \Omega(\bar{P}_N P_N\bar{z}))\to0$, $\Psi(\eta, \bar{\Omega})\to0$ при $\eta=(\delta, 1/N)\to0$. Используя их и переходя к пределу при $\eta\to0$ в (2), получаем искомую сходимость $\hat{\lambda}_\eta\to J^*$.

Лемма 1 представляет собой аналог теоремы 1.3.1, но непосредственно из нее не следует. Введем теперь множество $\hat{Z}_0\equiv\{\hat{z}\in\hat{D}:\ \hat{\Omega}(\hat{z})=\hat{\Omega}^*\}$.

Из результатов леммы 1.5.12, примененной к функционалам $\hat{I}$, $\hat{\Omega}$ с учетом лемм 1.1, 1.2, следует, что множество $\hat{Z}_0$ непусто и существуют такие элементы $\hat{z}_0 \in \hat{Z}_0$, для которых

$$\hat{I}(\hat{z}_0) = \inf\{\hat{I}(\hat{z}): \hat{z} \in \hat{Z}_0\} \equiv \hat{\nu}_\eta \equiv \hat{\nu}.$$

Последняя экстремальная задача является аналогом задачи (1.5.34).

Определим теперь по обычной схеме вспомогательные функции

$$\hat{\gamma}(\alpha) = \hat{\Omega}(\hat{z}^\alpha), \quad \hat{\beta}(\alpha) = \hat{I}(\hat{z}^\alpha), \quad \hat{\varphi}(\alpha) = \hat{M}^\alpha[\hat{z}^\alpha] \quad \forall \hat{z}^\alpha \in \hat{Z}^\alpha$$

(ср. § 1.6). Согласно результатам из § 1.6 и леммам 1.1, 1.2, эти функции обладают стандартными свойствами, перечисленными в леммах 1.6.1—1.6.6 и их следствиях. При этом фигурирующие в § 1.6 величины J_δ^*, Ω^*, ν_δ заменяются на их конечномерные аналоги $\hat{\mu}$, $\hat{\Omega}^*$, $\hat{\nu}$. Будем предполагать, что вместо числа $\hat{\lambda}$ известна его оценка сверху с точностью $\varkappa$, т. е. такое число $\hat{\lambda}_\varkappa \equiv \hat{\lambda}_{\eta\varkappa}$, для которого выполнено неравенство $0 \leqslant \hat{\lambda}_\varkappa - \hat{\lambda} \leqslant \varkappa$. Тогда можно ввести аналоги функций (1.7.1), (1.7.2):

$$\hat{\pi}(\alpha) = f[\hat{\lambda}_\varkappa + \Psi(\eta, \hat{\Omega}(\hat{z}^\alpha))] \equiv \hat{\Pi}[\hat{z}^\alpha],$$
$$\hat{\rho}(\alpha) = \hat{\beta}(\alpha) - \hat{\pi}(\alpha) \equiv \hat{P}[\hat{z}^\alpha].$$

Ясно, что эти функции обладают свойствами, перечисленными в теореме 1.7.1 с заменой чисел J_δ^*, λ_δ, ν_δ, Ω^* на их конечномерные аналоги.

Обратимся теперь к конструкции алгоритма *конечномерного обобщенного принципа невязки* (к.о.п.н.).

Выбор параметра регуляризации α_ζ $(\zeta \equiv (\eta, \varkappa))$ по алгоритму к.о.п.н. заключается в нахождении решения уравнения с монотонной функцией

$$\hat{\rho}(\alpha) = 0. \tag{3}$$

Из леммы 1.6.7 и теорем 1.7.2, 1.7.4 вытекает

Теорема 1. *Если для фиксированного ζ выполнено неравенство $\hat{\rho}_\infty \equiv \hat{\rho}(+\infty) > 0$, то уравнение* (3) *имеет по крайней мере одно решение $\alpha_\zeta \geqslant 0$. Если, кроме того, $\hat{\rho}_0 \equiv \hat{\rho}(+0) < 0$, то любое решение уравнения* (3) *положительно. Если при выполнении неравенств $\hat{\rho}_0 < 0$, $\hat{\rho}_\infty > 0$ задача минимизации сглаживающего функционала* (1.5) *имеет при каждом $\alpha > 0$ единственное решение $\hat{z}^\alpha$, причем для любых α_1, α_2 $(0 < \alpha_1 < \alpha_2)$ выполнено требование $\hat{z}^{\alpha_1} \neq \hat{z}^{\alpha_2}$, то уравнение* (3) *имеет единственное обычное решение $\alpha_\zeta > 0$.*

Приведем достаточные условия выполнения неравенств $\hat{\rho}_0 \equiv \hat{\rho}_0^\zeta < 0$, $\hat{\rho}_\infty \equiv \hat{\rho}_\infty^\zeta > 0$.

Теорема 2. *Пусть кроме условий постановки задачи из* § 1 *справедливы дополнительные условия*: а) $Z^* \cap Z_0 = \varnothing$; б) *хотя бы одна из функций $\Psi_0(\delta, \Omega)$, $\hat{\Psi}(\eta, \Omega)$ возрастает по второму аргументу при фиксированных δ $(\|\delta\| \neq 0)$, N. Тогда при достаточно малых $\|\eta\|$, $\varkappa$ $(0 < \|\eta\| \leqslant \Delta_0,\ 0 \leqslant \varkappa \leqslant \varkappa_0)$ выполнены неравенства $\hat{\rho}_0^\zeta < 0$, $\hat{\rho}_\infty^\zeta > 0$.*

Доказательство. Оно в основном следует схеме, использованной в теореме 1.7.3, так что подробно будут описаны лишь отличающиеся моменты доказательства. Это относится прежде всего к установлению сходимости $\hat{\nu}_\eta \to \nu_0$ при $\eta \to 0$.

Пусть $z_0 \in Z_0$ найдено, как в теореме 1.7.3, т. е.

$$\nu_0 = \inf\{f[J(z)]:\ z \in Z_0\} = f[J(z_0)],$$

а элемент $\hat{z}_0 \in \hat{Z}_0$ удовлетворяет условию

$$\hat{\nu} = \inf\{f[\hat{J}_\eta(\hat{z})]:\ \hat{z} \in \hat{Z}_0\} = f[\hat{J}_\eta(\hat{z}_0)].$$

Тогда в силу (1.8), (1.6)

$$\nu_0 = f[J(z_0)] \leqslant f[J(\bar{P}\hat{z}_0)] \leqslant f[\hat{J}_\eta(\hat{z}_0) + \Psi(\eta, \hat{\Omega}(\hat{z}_0))] = f[\hat{J}_\eta(\hat{z}_0) + \Psi(\eta, \hat{\Omega}^*)],$$

$$\hat{\nu} = f[\hat{J}_\eta(\hat{z}_0)] \leqslant f[\hat{J}_\eta(Pz_0)] \leqslant f[J(z_0) + \Psi(\eta, \Omega(z_0))] = f[J(z_0) + \Psi(\eta, \Omega^*)].$$

Из этих неравенств по аналогии с (1.7.5) получим

$$f[f^{-1}(\nu_0) - \Psi(\eta, \hat{\Omega}^*)] \leqslant \hat{\nu}_\eta \leqslant f[f^{-1}(\nu_0) + \Psi(\eta, \Omega^*)]. \tag{4}$$

По определению величины $\hat{\Omega}^*$ справедливо соотношение $\hat{\Omega}^*_N \leqslant \hat{\Omega}(P_N\bar{z}) = \Omega(\bar{P}_N P_N \bar{z})$, так что в силу условия 4) операторов P, $\bar{P}$ величины $\hat{\Omega}^*_N$ ограничены. Поэтому вследствие монотонности функции $\Psi(\eta, \Omega)$ по второму аргументу и в силу сходимости $\Psi(\eta, \Omega) \to 0$ при $\eta \to 0$ для каждого фиксированного Ω верны сходимости $\Psi(\eta, \hat{\Omega}^*_N) \to 0$, $\Psi(\eta, \Omega^*) \to 0$. Тогда из (4) с учетом непрерывности функции f получается сходимость $\hat{\nu}_\eta \to \nu_0$ при $\eta \to 0$.

Из этой сходимости, леммы 1 и теоремы 1.7.1 ясно, что

$$\lim_{\zeta \to 0} \hat{\rho}^\zeta_\infty = \lim_{\eta \to 0} \hat{\nu}_\eta - \lim_{\eta, \varkappa \to 0} f[\Psi(\eta, \hat{\Omega}^*_N) + \hat{\lambda}_\varkappa] = \nu_0 - \lim_{\eta, \varkappa \to 0} f[\hat{\lambda}_\varkappa] = \nu_0 - f(J^*) \equiv \nu_0 - \mu_0 > 0.$$

Здесь учтено доказанное в теореме 1.7.3 неравенство $\nu_0 > \mu_0$, вытекающее из условия $Z^* \bigcap Z_0 = \varnothing$. Полученное таким образом неравенство

$$\lim_{\zeta \to 0} \hat{\rho}^\zeta_\infty > 0$$

позволяет, как в теореме 1.7.3, установить, что $\hat{\rho}^\zeta_\infty > 0$ при $0 < \|\eta\| \leqslant \Delta_0$, $\varkappa \leqslant \varkappa_0$.

Доказательство неравенства $\hat{\rho}^\zeta_0 > 0$ проводится, как в теореме 1.7.3, с формальной заменой используемых там величин Ω_δ, Ω^*, λ_δ, δ на их аналоги $\hat{\Omega}_\eta \equiv \hat{\gamma}(+0)$, $\hat{\Omega}^*$, $\hat{\lambda}_\varkappa$, ζ. При этом используется вытекающее из условий теоремы монотонное возрастание функции $\Psi(\eta, \Omega)$ по второму аргументу.

Пусть теперь $q > 1$, $C > 1$ — заданные константы алгоритма и $\alpha_\zeta > 0$ определено по алгоритму к.о.п.н. Обозначим $\alpha_1 \equiv \alpha_\zeta / q$, $\alpha_2 \equiv \alpha_\zeta q$ и рассмотрим произвольные решения задачи (1.5) $\hat{z}^{\alpha_1}$, $\hat{z}^{\alpha_2}$, отвечающие соответственно значениям $\alpha = \alpha_{1,2}$.

Сформулируем правило отбора к.о.п.н. Если при фиксированном $\zeta=(\delta, 1/N, \varkappa)$ выполнено неравенство

$$\hat{I}(\hat{z}^{\alpha_2})\geqslant C\hat{\Pi}(\hat{z}^{\alpha_1})-(C-1)f(\hat{\lambda}_\varkappa), \tag{5}$$

то в качестве решения основной конечномерной задачи, отвечающего набору данных $(\hat{J}_\eta, \Psi_0, \hat{\Psi}, \delta, N, \varkappa)$, примем элемент $z_\zeta=\bar{P}_N\hat{z}_\zeta$, где $\hat{z}_\zeta\equiv\hat{z}^{\alpha_\zeta}$ выбирается из множества Z^{α_ζ} так, чтобы выполнялось неравенство

$$\hat{P}\left[\hat{z}^{\alpha_\zeta}\right]=\hat{I}(\hat{z}^{\alpha_\zeta})-\hat{\Pi}(\hat{z}^{\alpha_\zeta})\leqslant 0. \tag{6}$$

Например, можно взять $\hat{z}_\zeta=\hat{z}^{\alpha_\zeta}_{-}$ (см. лемму 1.6.5).

Если же оказывается, что

$$\hat{I}(\hat{z}^{\alpha_2})\leqslant C\hat{\Pi}(\hat{z}^{\alpha_1})-(C-1)f(\hat{\lambda}_\varkappa), \tag{7}$$

то решением основной конечномерной задачи считаем элемент $z_\zeta=\bar{P}_N\hat{z}^{\alpha_\zeta}$, где $\hat{z}^{\alpha_\zeta}$ выбрано из множества $\hat{Z}^{\alpha_\zeta}$ с таким расчетом, что выполнено неравенство

$$\hat{P}\left[\hat{z}^{\alpha_\zeta}\right]\geqslant 0. \tag{8}$$

В качестве $\hat{z}^{\alpha_\zeta}$ в этом случае можно принять, например, элемент $\hat{z}^{\alpha_\zeta}_{+}$.

Рассмотрим вопрос о сходимости приближенных решений.

Теорема 3. *Предположим, что для произвольной последовательности* $\{\zeta_n\}$, *где* $\zeta_n=(\delta_n, 1/N_n, \varkappa_n)$, *сходящейся к нулю при* $n\to\infty$, *определена, согласно алгоритму к.о.п.н., последовательность* $\{z_n\}$, *где* $z_n\equiv z_{\zeta_n}=\bar{P}_{N_n}\hat{z}^{\alpha_{\zeta_n}}$. *Тогда* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Доказательство. Убедимся, что для семейств $\{\bar{P}\hat{z}^{\alpha_\zeta}\}$, $\{\bar{P}\hat{z}^{\alpha_{1,2}}\}$ выполнены условия а)—д), приведенные в § 1.5. При этом роль чисел $\alpha(\delta)$, λ_δ из § 1.5 будут выполнять соответственно величины α_ζ, $\hat{\lambda}_\varkappa$. Введем также функционалы $J_\eta(\bar{P}\hat{z})\equiv\hat{J}_\eta(P\bar{P}\hat{z})=\hat{J}_\eta(\hat{z})$, $I_\eta(\bar{P}\hat{z})\equiv\equiv f[J_\eta(\bar{P}\hat{z})]=f[\hat{J}_\eta(\hat{z})]=\hat{I}(\hat{z})$, определенные при любом $\hat{z}\in\hat{D}$. Из (1.7) можно видеть, что выполнено неравенство

$$|J^*-J_\eta(\bar{P}\hat{\bar{z}})|\leqslant\Psi(\eta, \bar{\Omega}). \tag{9}$$

Выведем условие а). Для этого используем экстремальность элемента $\hat{z}^{\alpha_\zeta}$ как решения задачи (1.5), неравенство (9) и возрастание функции f:

$$\begin{aligned}\hat{M}^{\alpha_\zeta}[\hat{z}^{\alpha_\zeta}]&=\alpha_\zeta\hat{\Omega}(\hat{z}^{\alpha_\zeta})+\hat{I}(\hat{z}^{\alpha_\zeta})=\alpha_\zeta\Omega(\bar{P}\hat{z}^{\alpha_\zeta})+I_\eta(\bar{P}\hat{z}^{\alpha_\zeta})\leqslant\\&\leqslant\hat{M}^{\alpha_\zeta}[\bar{P}\bar{z}]=\alpha_\zeta\Omega(\bar{P}P\bar{z})+f[J_\eta(\bar{P}\hat{\bar{z}})]\leqslant\\&\leqslant\alpha_\zeta\Omega(\bar{P}P\bar{z})+f[J^*+\Psi(\eta, \bar{\Omega})]\leqslant\alpha_\zeta\bar{\Omega}_\eta+f[J^*+\Psi(\eta, \bar{\Omega}_\eta)].\end{aligned} \tag{10}$$

Здесь принято обозначение $\bar{\Omega}_\eta\equiv\max\{\bar{\Omega}, \Omega(\bar{P}_NP_N\bar{z})\}$ $(\bar{\Omega}_\eta\geqslant\bar{\Omega})$. Неравенство (10) представляет собой условие а) с $\bar{\Omega}_\delta=\bar{\Omega}_\eta$, так как вследствие условия 4) операторов $\bar{P}_N$, P_N имеет место сходимость $\bar{\Omega}_\eta\to\bar{\Omega}$ при $\eta=(\delta, 1/N)\to 0$.

Аналогичным образом устанавливается и условие б), имеющее в данном случае вид

$$\hat{M}^{\alpha_2}[\hat{z}^{\alpha_2}]=\alpha_2\Omega(\bar{P}\hat{z}^{\alpha_2})+I_\eta(\bar{P}\hat{z}^{\alpha_2})\leqslant\alpha_\zeta q\bar{\Omega}_\eta+f[J^*+\Psi(\eta, \bar{\Omega}_\eta)].$$

Условие в), связанное с монотонностью вспомогательных функций $\hat{\gamma}$, $\hat{\beta}$, следует из леммы 1.6.1 и неравенства $\alpha_1<\alpha_\zeta<\alpha_2$:

$$\Omega(\bar{P}\hat{z}^{\alpha_1})=\hat{\Omega}(z^{\alpha_1})=\hat{\gamma}(\alpha_1)\geqslant\hat{\gamma}(\alpha_\zeta)=\Omega(\bar{P}\hat{z}^{\alpha_\zeta})\geqslant\hat{\gamma}(\alpha_2)=\Omega(\bar{P}\hat{z}^{\alpha_2}),$$
$$I_\eta(\bar{P}\hat{z}^{\alpha_1})=\hat{I}(\hat{z}^{\alpha_1})=\hat{\beta}(\alpha_1)\leqslant\hat{\beta}(\alpha_\zeta)=I_\eta(\bar{P}\hat{z}^{\alpha_\zeta})\leqslant\hat{\beta}(\alpha_2)=I_\eta(\bar{P}\hat{z}^{\alpha_2}).$$

Так же обосновывается и условие д).

Наконец, условие г) выполнено в силу принятого правила отбора (5)—(8).

Выполнение условий а)—д) влечет справедливость для величин $\Omega(\bar{P}\hat{z}^{\alpha_\zeta})$, $I_\eta(\bar{P}\hat{z}^{\alpha_\zeta})$ лемм 1.5.7—1.5.10. Из них в свою очередь, как отмечено в § 1.5, следует, что

$$\varlimsup_{\zeta\to0}\Omega(\bar{P}_N\hat{z}^{\alpha_\zeta})\leqslant\bar{\Omega},\quad \lim_{\zeta\to0}I_\eta(P_N\hat{z}^{\alpha_\zeta})=I^*=f(J^*).$$

Эти предельные соотношения по лемме 1.5.1 и следствию 1.5.1 обеспечивают доказываемые сходимости: $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ при $n\to\infty$.

Доказанная теорема показывает, что алгоритм к.о.п.н. дает решение основной конечномерной задачи и является регуляризующим алгоритмом для решения вариационной задачи (1.1).

§ 3. Конечномерные обобщенные принципы квазирешений и сглаживающего функционала

Конечномерный обобщенный принцип квазирешений (к.о.п.к.) для своей реализации использует кроме данных $\{\hat{J}_\eta,\ \Psi_0,\ \hat{\Psi},\ \delta,\ N,\ Z_N\}$ величину $\hat{\omega}_\eta$, представляющую собой оценку для $\bar{\Omega}$.

Выбор параметра регуляризации α_ζ по алгоритму к.о.п.к. заключается в нахождении решения уравнения с монотонной функцией

$$\hat{\gamma}(\alpha)=\hat{\omega}_\eta. \tag{1}$$

Из леммы 1.6.7 и теорем 1.8.1, 1.8.2 следует

Теорема 1. *Если для фиксированного* η *выполнено неравенство* $\hat{\Omega}^*<\hat{\omega}_\eta<\hat{\Omega}_\eta$, *то уравнение* (1) *имеет по крайней мере одно решение* $\alpha_\eta>0$. *Если, кроме того, экстремальная задача* (1.5) *имеет при каждом* $\alpha>0$ *единственное решение* $\hat{z}^\alpha$, *причем* $\hat{z}^{\alpha_1}\neq\hat{z}^{\alpha_2}$ *для любых* α_1, α_2 $(0<\alpha_1<\alpha_2)$, *то уравнение* (1) *имеет единственное обычное решение.*

Предположим теперь, что $\alpha_\eta>0$ найдено как решение уравнения (1). Сформулируем правило отбора конечномерного о.п.к. Пусть $C>1$, $q>1$ — константы $\alpha_1\equiv\alpha_\eta/q$, $\alpha_2\equiv\alpha_\eta q$. Пусть также $\hat{z}^{\alpha_1}$, $\hat{z}^{\alpha_2}$ — произвольные решения задачи (1.5) соответственно для $\alpha=\alpha_1$, α_2.

Если выполнено неравенство (2.5), то выберем из множества $\hat{Z}^{\alpha_\eta}$ такой элемент $\hat{z}^{\alpha_\eta}$, для которого $\hat{\Omega}(\hat{z}^{\alpha_\eta})\geqslant\hat{\omega}_\eta$. Например, можно взять $\hat{z}^{\alpha_\eta}=\hat{z}_-^{\alpha_\eta}$.

Если же выполнено условие (2.7), то элемент $\hat{z}^{\alpha_\eta}$ выбирается из $\hat{Z}^{\alpha_\eta}$ так, чтобы было справедливо условие $\hat{\Omega}(\hat{z}^{\alpha_\eta})\leqslant\hat{\omega}_\eta$. Например,

$\hat{z}^{\alpha_\eta}=\hat{z}_+^{\alpha_\eta}$. Отметим, что выбранный таким образом элемент $\hat{z}^{\alpha_\eta}$ будет зависеть и от числа $\hat{\lambda}_\varkappa$ (см. (2.5), (2.7)), т. е. и от $\varkappa$: $\hat{z}^{\alpha_\eta}=\hat{z}^{\alpha_\eta}(\varkappa)$.

В качестве решения основной конечномерной задачи используем величины $z_\zeta=\overline{P}_N\hat{z}^{\alpha_\eta}(\varkappa)$.

Теорема 2. *Пусть числа $\hat{\omega}_\eta$ подчиняются условиям $\hat{\omega}_\eta\geqslant\overline{\Omega}$, $\hat{\omega}_\eta\to\overline{\Omega}$ при $\eta\to 0$. Предположим также, что для последовательности $\{\zeta_n\}$, сходящейся к нулю при $n\to\infty$, величины $\alpha_{\eta_n}>0$, $\hat{z}^{\alpha_{\eta_n}}(\varkappa_n)$ найдены по алгоритму к.о.п.к. Тогда для последовательности $\{z_n\}$, где $z_n=\overline{P}_{N_n}\hat{z}^{\alpha_{\eta_n}}(\varkappa_n)$, справедливы предельные соотношения $z_n\xrightarrow{\tau}\overline{Z}$, $\Omega(z_n)\to\overline{\Omega}$ при $n\to\infty$.*

Доказательство. Оно проводится по той же схеме, что и в теореме 2.3, и связано с применением для семейства $\{P_N\hat{z}^{\alpha_\eta}(\varkappa)\}$ леммы 1.5.11 и следствия 1.5.1. Проверка фигурирующих в условии леммы 1.5.11 требований а)—д) производится так же, как это делалось при доказательстве теоремы 2.3 с заменой чисел α_ζ на числа α_η. Модификация условий г), принятая в лемме 1.5.11, содержится в условиях правила отбора алгоритма к.о.п.к. Итак, по теореме 2 семейство $\{P_N\hat{z}^{\alpha_\eta}(\varkappa)\}$ дает решение основной конечномерной задачи.

Обратимся теперь к *конечномерному обобщенному принципу сглаживающего функционала* (к.о.п.с.ф.). Введем функцию

$$\hat{\varepsilon}(\alpha)\equiv\hat{\varphi}(\alpha)-f[\hat{\lambda}_\varkappa+\Psi^p(\eta,\ \hat{\Omega}(\hat{z}^\alpha))]\equiv\hat{E}(\hat{z}^\alpha),\quad \alpha>0,$$

где параметр p фиксирован и удовлетворяет условию $0<p<1$. Функция $\hat{\varepsilon}(\alpha)$ представляет собой аналог функции (1.9.1) и обладает стандартными свойствами, перечисленными в теореме 1.9.1 с заменой фигурирующих там величин Ω^*, λ_δ, ν_δ на $\hat{\Omega}^*$, $\hat{\lambda}_\varkappa$, $\hat{\nu}$. В дальнейшем, как и в § 1.9, будем считать, что $\Omega^*\geqslant 0$ и, следовательно, $\hat{\Omega}_N^*\geqslant 0$.

Выбор параметра регуляризации α_ζ по алгоритму к.о.п.с.ф. осуществляется путем нахождения решения уравнения с монотонной функцией

$$\hat{\varepsilon}(\alpha)=0. \tag{2}$$

Следуя теоремам 1.9.2—1.9.4, можно убедиться, что верны

Теорема 3. *Если при фиксированном ζ выполнены неравенства $\hat{\varepsilon}_0\equiv\hat{\varepsilon}(+0)<0$, $\hat{\varepsilon}_\infty\equiv\hat{\varepsilon}(+\infty)>0$, то уравнение* (2) *имеет единственное решение $\alpha_\zeta>0$. Это решение будет обычным, если задача* (1.5) *минимизации сглаживающего функционала имеет единственное решение $\hat{z}^\alpha$ при любом $\alpha>0$.*

Теорема 4. *Пусть выполнены условия теоремы* 2.2. *Тогда при достаточно малых $\|\eta\|$, $\varkappa$ выполнены неравенства $\hat{\varepsilon}_0<0$, $\hat{\varepsilon}_\infty>0$.*

Доказательство. Оно проводится, как в теореме 2.2, с заменой величин $\hat{\rho}_0$, $\hat{\rho}_\infty$ на $\hat{\varepsilon}_0$, $\hat{\varepsilon}_\infty$ и с введением вместо функции $\Psi(\eta,\ \Omega)$ функции $\Psi^p(\eta,\ \Omega)$ (см. также теорему 1.9.3).

Предположим теперь, что параметр регуляризации $\alpha_\zeta>0$ выбран как решение уравнения (2). Сформулируем правило отбора алгоритма к.о.п.с.ф. В качестве решения основной конечномерной задачи примем величину $z_\zeta\equiv\overline{P}_N\hat{z}^{\alpha_\zeta}$, где элемент $\hat{z}^{\alpha_\zeta}$ выбран из множества

$\hat{Z}^{\varkappa}$ так, чтобы выполнялось неравенство $\hat{E}(\hat{z}^{\varkappa})\geqslant 0$. Так, например, можно взять $\hat{z}^{\varkappa}$ равным элементу $\hat{z}^{\varkappa}_{\dagger}$.

Теорема 5. *Пусть в дополнение к условиям постановки задачи § 1 выполнено условие $Z^*\bigcap Z_0=\varnothing$. Тогда $\alpha_\zeta\to 0$ при $\zeta\to 0$.*

Доказательство. Оно проводится, как в теореме 1.9.5, если заменить там α_δ на α_ζ, а вместо z^α взять величины $\bar{P}\hat{z}^\alpha$. Кроме того, там формально необходимо заменить λ_δ, Ω^*, $\Psi(\delta,\ \Omega)$ соответственно на их аналоги $\hat{\lambda}_\varkappa$, $\hat{\Omega}^*$, $\Psi(\eta,\ \Omega)$.

Как и в теореме 1.9.6, при замене соответствующих обозначений и с использованием в качестве чисел $\bar{\Omega}_\delta$ величин $\bar{\Omega}_\eta\equiv\max\{\bar{\Omega},\ \Omega(\bar{P}P\bar{z})\}$ может быть доказана следующая

Теорема 6. *Если выполнены условия теоремы 2.2 и величины $z_\zeta\equiv\bar{P}\hat{z}^{\alpha_\zeta}$ найдены по алгоритму к.о.п.с.ф., то для любой последовательности $\{\zeta_n\}$, сходящейся к нулю при $n\to\infty$, соответствующая последовательность $\{z_n\}$, где $z_n\equiv z_{\zeta_n}$, обладает свойствами $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ при $n\to\infty$.*

Итак, алгоритм к.о.п.с.ф. также дает решение основной конечномерной задачи.

§ 4. Конечномерные алгоритмы для операторных уравнений

Алгоритмы решения некорректных экстремальных задач, изложенные в § 2, 3 и использующие для построения приближенного решения конечномерные данные, могут быть использованы и для решения вариационным методом операторных уравнений. Дадим соответствующую постановку задачи.

Пусть $(Z,\ \tau)$ — хаусдорфово топологическое пространство, U — метрическое пространство с метрикой ρ. Предположим, что на некотором множестве D $(D\subset Z)$ определен в общем случае нелинейный оператор $A\colon D\to U$, и пусть $\bar{u}\in U$ — некоторый элемент. Введем операторное уравнение: для величин $(A,\ \bar{u})$ найти элементы $z\in D$ такие, что

$$Az=\bar{u}. \tag{1}$$

Предположим, что уравнение (1) имеет непустое множество Z^* квазирешений на D (см. § 2.1). Определим на множестве D функционал $\Omega(z)$ $(\Omega(z)\geqslant\Omega^*)$ и будем рассматривать задачу о поиске Ω-оптимальных квазирешений уравнения (1) на множестве D: найти элементы $\bar{z}\in Z^*$ такие, для которых

$$\Omega(\bar{z})=\inf\{\Omega(z)\colon z\in Z^*\}. \tag{2}$$

Как обычно, будем обозначать множество решений задачи (2) в виде $\bar{Z}$.

В § 2.1 рассмотрена основная задача нахождения приближений к множеству $\bar{Z}$ по приближенным данным задачи. При этом считалось, что вместо величин $(A,\ \bar{u})$ имеются их приближения $(A_h,\ u_\sigma)$, где оператор A_h действует из D в U и удовлетворяет условию аппроксимации

$$\rho(Az,\ A_h z)\leqslant\psi(h,\ \Omega(z))\quad\forall z\in D,$$

а элемент $u_\sigma\in U$ подчинен условию $\rho(\bar{u},\ u_\sigma)\leqslant\sigma$. Числа h, σ, а также функция ψ предполагаются известными.

В данном параграфе мы будем считать, что величины (A_h, u_σ) также не известны, как и $(\bar{A}, \bar{u})$, а в нашем распоряжении имеются их конечномерные приближения. Перейдем к постановке задачи конечномерной аппроксимации [134].

Зададим конечномерные нормированные пространства Z_N, U_M соответственно размерности N, M. Будем обозначать элементы этих пространств символами $\hat{z}$, $\hat{u}$: $\hat{z}\in Z_N$, $\hat{u}\in U_M$. Как и в § 1, введем операторы $P_N\colon Z\to Z_N$, $\bar{P}_N\colon Z_N\to Z$, обладающие перечисленными там свойствами. Тогда можно ввести непустое множество $\hat{D}=\{\hat{z}\colon \bar{P}\hat{z}\in D\}$ — конечномерный аналог множества D. Определим также оператор $\bar{Q}_M$, непрерывный из U_M в U. Будем считать, что вместо приближенных данных (A_h, u_σ) основной задачи известны их конечномерные приближения, т. е. оператор $\hat{A}_H$, где $H\equiv(h, 1/M, 1/N)$, действующий из $\hat{D}$ в U_M, а также элемент $\hat{u}_\sigma\equiv\hat{u}_{\sigma M}\in U_M$. При этом выполнены условия конечномерной аппроксимации:

$$\rho(u_\sigma, \bar{Q}_M\hat{u}_\sigma)\leqslant\xi(1/M, \sigma), \tag{3}$$

$$\rho(A_h z, \bar{Q}_M\hat{A}_H P_N z)\leqslant\hat{\psi}(H, \Omega(z)) \quad \forall z\in D, \tag{4}$$

где $\xi(1/M, \sigma)$, $\hat{\psi}(H, \Omega(z))$ — известные меры аппроксимации.

Основная конечномерная задача, которая и будет изучаться в данном параграфе, заключается в нахождении по набору приближенных данных $\{\hat{A}_H, \hat{u}_\sigma, \hat{\psi}, \xi, \psi, h, \delta, M, N, Z_N, U_M\}$ такого элемента $\hat{z}_\eta\in\hat{D}$ $(\eta=(H, \sigma))$, что семейство $\{z_\eta\}$, где $z_\eta\equiv\bar{P}_N\hat{z}_\eta$, будет τ-секвенциально сходиться к $\bar{Z}$ при $\eta\to 0$.

Вводя функционалы

$$J(z)\equiv\rho(Az, \bar{u}), \quad J_\delta(z)\equiv\rho(A_h z, u_\sigma),$$

$$\Psi_0(\delta, \Omega(z))\equiv\sigma+\psi(h, \Omega(z)), \quad z\in D, \quad \delta\equiv(h, \sigma),$$

$$\hat{J}_\eta(\hat{z})\equiv\rho(\bar{Q}_M\hat{A}_H\hat{z}, \bar{Q}_M\hat{u}_\sigma), \quad \hat{I}(\hat{z})\equiv f[\hat{J}_\eta(\hat{z})],$$

$$\hat{\Omega}(\hat{z})\equiv\Omega(\bar{P}\hat{z}), \quad \hat{\Psi}(\eta, \Omega(z))\equiv\xi(1/M, \sigma)+\hat{\psi}(H, \Omega(z)),$$

$$\Psi(\eta, \Omega)\equiv\Psi_0(\delta, \Omega)+\hat{\Psi}(\eta, \Omega)=\sigma+\xi(1/M, \sigma)+\psi(h, \Omega)+\hat{\psi}(H, \Omega),$$

легко убедиться, что эта основная конечномерная задача представляет собой частный случай более общей основной конечномерной задачи для некорректных экстремальных задач из § 1.

В дальнейшем будем считать, что кроме уже перечисленных требований постановки задачи выполнены предположения 1—4 из § 2.1, свойства 1)—4) операторов $\bar{P}$, P из § 1, а также следующие дополнительные условия:

а) операторы $\hat{A}_H$ непрерывны из $\hat{D}$ в U_M для каждого фиксированного H;

б) функции $\xi(s, \sigma)$, определенные при s, $\sigma\geqslant 0$, непрерывны и подчинены равенству $\xi(0, 0)=0$;

в) мера аппроксимации $\hat{\psi}(H, \Omega)$ удовлетворяет стандартным требованиям, предъявленным к мере аппроксимации в § 1.2.

Выполнение всех этих требований гарантирует применимость результатов § 1—3 для решения рассматриваемой основной конечномерной задачи. В частности, для нахождения приближенных решений

$z_\eta = \bar{P}_N \hat{z}_\eta$ можно использовать конечномерные обобщенные принципы невязки, квазирешений и сглаживающего функционала в форме, данной в § 2, 3.

Остановимся на частном случае задачи (1), когда это операторное уравнение разрешимо на D. Тогда, как и в § 2.3, можно построить модификации конечномерных алгоритмов, не использующие величины обобщенной конечномерной меры несовместности:

$$\hat{\lambda} = \inf\{\hat{J}_\eta(\hat{z}) + \Psi(\eta, \hat{\Omega}(z)) \colon \hat{z} \in \hat{D}\} = \inf\{\rho(\hat{Q}\hat{A}_H \hat{z}, \hat{u}_\sigma) + \\ + \sigma + \xi(1/M, \sigma) + \psi(h, \hat{\Omega}(\hat{z})) + \hat{\psi}(H, \hat{\Omega}(\hat{z})) \colon \hat{z} \in \hat{D}\}$$

или ее приближенных значений. Итак, пусть уравнение (1) разрешимо на множестве D, т. е. $\mu_0 \equiv J^* = 0$. Сформулируем алгоритм *модифицированного конечномерного о.п.н.* (м.к.о.п.н.). Для этого определим функции

$$\hat{\bar{\pi}}(\alpha) \equiv Kf[\Psi(\eta, \hat{\Omega}(\hat{z}^\alpha))] = K\hat{\bar{\Pi}}(\hat{z}^\alpha),$$
$$\hat{\bar{\rho}}(\alpha) \equiv \hat{\beta}(\alpha) - \hat{\bar{\pi}}(\alpha) = \hat{\bar{P}}(\hat{z}^\alpha), \quad \hat{z}^\alpha \in \hat{Z}^\alpha, \quad K = \text{const} \geqslant 1.$$

Выбор параметра регуляризации α_η по алгоритму м.к.о.п.н. заключается в нахождении решения уравнения с монотонно не убывающей функцией

$$\hat{\bar{\rho}}(\alpha) = 0, \quad \alpha \geqslant 0. \tag{5}$$

Существование такого решения α_η и условия его положительности обоснованы в теоремах 2.1, 2.2.

Элемент $\hat{z}_\eta \equiv \hat{z}^{\alpha_\eta}$, используемый для решения основной конечномерной задачи данного параграфа, находится по правилу отбора алгоритма м.к.о.п.н. При этом, как и в § 2, используются величины C, $q > 1$, $\alpha_1 = \alpha_\eta / q$, $\alpha_2 = \alpha_\eta q$, $\hat{z}^{\alpha_1}$, $\hat{z}^{\alpha_2}$.

а) Пусть $\alpha_\eta > 0$. Если выполнено неравенство

$$\hat{I}(\hat{z}^{\alpha_2}) \geqslant C \hat{\bar{\Pi}}(\hat{z}^{\alpha_1}), \tag{6}$$

то элемент $\hat{z}^{\alpha_\eta}$ выбирается из множества $\hat{Z}^{\alpha_\eta}$ так, чтобы выполнялось условие $\bar{P}(\hat{z}^{\alpha_\eta}) \leqslant 0$. Например, можно считать, что $\hat{z}^{\alpha_\eta} = \hat{z}^{\alpha_\eta}_-$. Если же оказывается, что

$$\hat{I}(\hat{z}^{\alpha_2}) \leqslant C \hat{\bar{\Pi}}(\hat{z}^{\alpha_1}), \tag{7}$$

то выбирается такой элемент $\hat{z}^{\alpha_\eta} \in \hat{Z}^{\alpha_\eta}$, для которого $\hat{\bar{P}}(\hat{z}^{\alpha_\eta}) \geqslant 0$. В частности, можно принять $\hat{z}^{\alpha_\eta} = \hat{z}^{\alpha_\eta}_+$.

б) Если $\alpha_\eta = 0$, то элемент $\hat{z}_\eta$ выбирается как любое решение экстремальной задачи: найти такие $\hat{z}_\eta \in \hat{D}$, для которых

$$\hat{J}_\eta(\hat{z}_\eta) = \inf\{\hat{J}_\eta(\hat{z}) \colon \hat{z} \in \hat{Z}_0\}. \tag{8}$$

Задача (8) разрешима, как отмечено в § 2.

Рассмотрим вопрос о сходимости приближений $z_\eta = \bar{P}_N \hat{z}_\eta$.

Теорема 1. *Пусть $\{\eta_n\}$ — произвольная последовательность, сходящаяся к нулю при $n \to \infty$, и $\alpha_{\eta_n} > 0$ для каждого n. Тогда при выполнении перечисленных выше условий и при дополнительном условии непрерывности и монотонного возрастания функции $f(x)$ при $x \geqslant 0$ построенная с помощью алгоритма м.к.о.п.н. последовательность $\{z_n\}$, где $z_n \equiv \bar{P}_{N_n} \hat{z}^{\alpha_{\eta_n}}$, обладает свойствами $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ при $n \to \infty$.*

Доказательство проводится абсолютным повторением доказательства теоремы 2.3.1, основано на использовании аналогов лемм 2.3.1—2.3.3 и отличается лишь обозначениями входящих в соответствующие формулы величин. В частности, в аналогах формул (2.3.8)—(2.3.13) следует вместо величины $\bar{\Omega}$ использовать число $\bar{\Omega}_\eta \equiv \max\{\bar{\Omega}, \Omega(\bar{P}_N P_N \bar{z})\}$, как это, например, делалось в теореме 2.3, с учетом того, что $\bar{\Omega}_\eta \to \bar{\Omega}$ при $\eta \to 0$.

Теорема 2. *Пусть в дополнение к условиям постановки задачи совокупная мера аппроксимации* $\Psi(\eta, \Omega)$ *возрастает по второму аргументу при каждом фиксированном ненулевом* η. *Если для последовательности* $\{\eta_n\}$, *где* $\eta_n \neq 0$, *сходящейся к нулю при* $n \to \infty$, *оказывается, что* $\alpha_{\eta_n} = 0$ *и элементы* $\hat{z}_n \equiv \hat{z}_{\eta_n}$ *получены как решения задачи* (8), *то последовательность* $\{z_n\}$, *где* $z_n \equiv \bar{P}_{N_n} \hat{z}_n$, *обладает свойствами* $z_n \xrightarrow{\tau} \bar{Z}$, $\Omega(z_n) \to \bar{\Omega}$ *при* $n \to \infty$.

Доказательство проводится по той же схеме, что и в теореме 1.10.3. Отметим специфические особенности. Прежде всего из принадлежности $\hat{z}_\eta \in \hat{Z}_0$ следует оценка

$$\Omega(\bar{P}\hat{z}_\eta) = \hat{\Omega}(\hat{z}_\eta) = \hat{\Omega}^* \leqslant \Omega(\bar{P}_N P_N \bar{z}),$$

откуда, используя свойство 4) операторов $\bar{P}$, P, получаем предельное соотношение

$$\overline{\lim_{\eta \to 0}} \Omega(\bar{P}_N \hat{z}_\eta) \leqslant \overline{\lim_{\eta \to 0}} \Omega(\bar{P}_N P_N \bar{z}) \leqslant \bar{\Omega}, \tag{9}$$

являющееся аналогом оценки (1.10.10). Далее из неравенств (1.7), (1.8) и аналога теоремы 1.10.1 следует с учетом того, что $J^* = 0$, оценка

$$0 - \Psi(\eta, \hat{\Omega}^*) \leqslant \hat{J}_\eta(\hat{z}_\eta) \equiv J_\eta(\bar{P}\hat{z}_\eta) = \inf\{\hat{J}_\eta(\hat{z}) : \hat{z} \in \hat{Z}_0\} =$$
$$= \inf\{\hat{J}_\eta(\hat{z}) : \hat{z} \in \hat{D}\} \leqslant \hat{J}_\eta(\hat{\bar{z}}) \leqslant 0 + \Psi(\eta, \bar{\Omega}),$$

из которой, как и в теореме 1.10.3, получим

$$\lim_{\eta \to 0} J_\eta(\bar{P}_N \hat{z}_\eta) = 0. \tag{10}$$

Из (9), (10) по лемме 1.5.1 получаются искомые сходимости.

Таким образом, алгоритм м.к.о.п.к. можно использовать для решения основной конечномерной задачи в случае разрешимого уравнения (1).

Рассмотрим теперь *модифицированный к.о.п.к.* (*м.к.о.п.к.*). Выбор параметра регуляризации α_η в этом алгоритме осуществляется из уравнения (3.1), а правило отбора элемента $\hat{z}^{\alpha_\eta}$ с учетом обозначений из § 3 имеет вид:

а) пусть $\alpha_\eta > 0$; если выполнено неравенство (6), то элемент $\hat{z}^{\alpha_\eta}$ выбирается из $\hat{Z}^{\alpha_\eta}$ так, чтобы было выполнено условие $\hat{\Omega}(\hat{z}^{\alpha_\eta}) \geqslant \hat{\omega}_\eta$; если же справедливо соотношение (7), то элемент $\hat{z}^{\alpha_\eta}$ отбирается из условия $\hat{\Omega}(\hat{z}^{\alpha_\eta}) \leqslant \hat{\omega}_\eta$;

б) если $\alpha_\eta = 0$, то в качестве $\hat{z}^{\alpha_\eta}$ берется любое решение экстремальной задачи: найти элементы $\hat{z}^0_\eta$, для которых

$$\hat{J}_\eta(\hat{z}^0_\eta) = \inf\{\hat{J}_\eta(\hat{z}) : \hat{z} \in D\}. \tag{11}$$

Разрешимость задачи (11) устанавливается, как и для ее аналога (1.10.5) в теореме 1.10.2. При этом оказывается, что $\hat{\Omega}(\hat{z}_\eta^0)\leqslant\hat{\omega}_\eta$.

Справедлива следующая теорема сходимости приближений $z_\eta=\bar{P}\hat{z}_\eta$.

Теорема 3. *Пусть выполнены условия постановки задачи и числа $\hat{\omega}_\eta$ подчинены ограничениям $\hat{\omega}_\eta\geqslant\bar{\Omega}$, $\hat{\omega}_\eta\to\bar{\Omega}$ при $\eta\to 0$. Пусть, кроме того, функция $f(x)$ непрерывна и монотонно возрастает при $x\geqslant 0$. Тогда для любой последовательности $\{\eta_n\}$, сходящейся к нулю при $n\to\infty$, соответствующая последовательность $\{z_n\}$, где $z_n\equiv z_{\eta_n}$, полученная с помощью алгоритма м.к.о.п.к., обладает свойствами $z_n\xrightarrow{\text{I}}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ при $n\to\infty$.*

Доказательство. В случае $\alpha_{\eta_n}>0$ оно проводится, как это указано в теореме 3.2. Если же $\alpha_{\eta_n}=0$ для всякого n, то из (11) и условия аппроксимации (1.7) получим

$$0\leqslant J_\eta(\bar{P}_N\hat{z}_\eta)\leqslant J_\eta(\bar{P}_N P_N\bar{z})\leqslant 0+\Psi(\eta,\bar{\Omega}),$$

откуда понятно, что $J_\eta(\bar{P}_N\hat{z}_\eta)\to 0$ при $\eta\to 0$. Кроме того, из свойства решений задачи (11)

$$\varlimsup_{\eta\to 0}\Omega(\bar{P}_N\hat{z}_\eta)\leqslant\lim_{\eta\to 0}\hat{\omega}_\eta=\bar{\Omega}.$$

Для получения искомых сходимостей остается применить лемму 1.5.1.

Обратимся, наконец, к *модифицированному к.о.п.с.ф.* (м.к.о.п.с.ф.). Введем функцию

$$\hat{\bar{\varepsilon}}(\alpha)\equiv\hat{\varphi}(\alpha)-f[\Psi^p(\eta,\hat{\Omega}(\hat{z}^\alpha))]\equiv\hat{\bar{E}}(\hat{z}^\alpha),$$

аналогичную функции $\bar{\varepsilon}(\alpha)$ из § 2.3.

Параметр регуляризации α_η, удовлетворяющий алгоритму м.к.о.п.с.ф., находится как решение уравнения с монотонной функцией

$$\hat{\bar{\varepsilon}}(\alpha)=0,\quad \Omega^*\geqslant 0. \tag{12}$$

Для решения этого уравнения справедлив результат теоремы 3.3. Как и в теореме 2.2, можно доказать, что если $Z_0\cap Z^*=\varnothing$, то при $\|\eta\|\leqslant\Delta_0$ уравнение (12) имеет единственное решение $\alpha_\eta>0$.

Правило отбора м.к.о.п.с.ф.:

а) если $\alpha_\eta>0$, то элемент $\hat{z}_\eta\equiv z^{\alpha_\eta}$ выбирается из множества $\hat{Z}^{\alpha_\eta}$ так, чтобы выполнялось неравенство $\bar{E}(\hat{z}^{\alpha_\eta})\geqslant 0$, например, $\hat{z}^{\alpha_\eta}=\hat{z}_+^{\alpha_\eta}$;

б) если $\alpha_\eta=0$, то элемент $\hat{z}_\eta$ выбирается как произвольное решение задачи (8).

В качестве решения основной конечномерной задачи берется величина $z_\eta=\bar{P}_N\hat{z}_\eta$.

Используя технику доказательства теорем 1.9.2—1.9.4 с учетом замечаний, данных в теоремах 3.4—3.6 и теореме 2, можно установить следующие теоремы.

Теорема 4. *При выполнении условий постановки задачи и дополнительного условия $Z^*\cap Z_0=\varnothing$ параметр регуляризации, выбранный по алгоритму м.к.о.п.с.ф., обладает свойством $\alpha_\eta\to 0$ при $\eta\to 0$.*

Теорема 5. *При выполнении условий постановки задачи и условий:* 1) $Z^*\bigcap Z_0=\varnothing$; 2) *совокупная мера аппроксимации* $\Psi(\eta, \Omega)$ *возрастает по второму аргументу при каждом* η, *отличном от нуля; для любой последовательности* $\{\eta_n\}$, *сходящейся к нулю при* $n\to\infty$, *соответствующая последовательность* $\{z_n\}$, *где* $z_n\equiv\bar{P}_{N_n}\hat{z}_{\eta_n}$, *такова, что* $z_n\xrightarrow{\tau}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ *при* $n\to\infty$.

Из результатов данного параграфа следует, что модифицированные конечномерные принципы невязки, квазирешений и сглаживающего функционала представляют собой регуляризующие алгоритмы для решения разрешимых операторных уравнений.

§ 5. Примеры задач

В данном параграфе будут рассмотрены часто встречающиеся на практике примеры операторных уравнений, для нахождения Ω-оптимальных решений которых можно применять изложенные в § 2—4 конечномерные алгоритмы. При рассмотрении этих примеров главный акцент будет делаться на конкретизацию общей схемы конечномерной аппроксимации некорректных задач, введенной в § 1. Для этого будем рассматривать случай $h=0$, $\sigma=0$ (см. § 4), в котором алгоритмы учитывают исключительно погрешность конечномерной аппроксимации данных задачи.

Пример 1. Предположим, что функция $K(x, s, z)$ определена на множестве $\Pi_0\equiv\{[c, d]\times[a, b]\times\mathbf{R}\}$ и удовлетворяет условиям: а) $K(x, s, z)\in C^1(\Pi_0)$; б) для любого $(x, s, z)\in\Pi_0$ справедливы оценки

$$|K(x, s, z)|\leqslant\xi_0(|z|),\quad |K'_x(x, s, z)|\leqslant\xi_1(|z|),$$

$$|K'_s(x, s, z)|\leqslant\xi_2(|z|),\quad |K'_z(x, s, z)|\leqslant\xi_3(|z|);$$

здесь функции $\xi_m(x)(m=0, 1, 2, 3)$ известны, непрерывны и монотонно не убывают при $x\geqslant 0$. Определим на множестве функций $z(s)\in C[a, b]$ оператор

$$A[x, z(s)]\equiv\int_a^b K(x, s, z(s))\,ds,\quad x\in[c, d]. \tag{1}$$

Этот оператор является усиленно непрерывным из $W_2^1[a, b]$ в $C[c, d]$ (или $L_2[c, d]$). Действительно, из неравенства

$$|K(x, s, z_1(s))-K(x, s, z_2(s))|\leqslant\xi_3(z^*)|z_1(s)-z_2(s)|,$$

которое справедливо для непрерывных функций $z_{1,2}(s)$ и которое вытекает из формулы конечных приращений и условия б) при $z^*\equiv\max(\|z_1\|_C, \|z_2\|_C)$, следует с учетом результата примера 2.4.4 непрерывность оператора (1) из пространства $C[a, b]$ в $C[c, d]$. Поэтому вследствие полной непрерывности оператора вложения пространства $W_2^1[a, b]$ в $C[a, b]$ (см. теорему 0.4.9) оператор (1) будет также усиленно непрерывным.

Пусть теперь пространство Z совпадает с $W_2^1[a, b]$, а топология τ в нем — топология слабой сходимости. Определим также

$U = L_2[c, d]$ с топологией t сходимости по норме. Введем операторное уравнение

$$A[x, z(s)] = \bar{u}(x), \quad x \in [c, d], \tag{2}$$

и будем предполагать, что оно имеет для правой части $\bar{u}(x) \in C^1[c, d] \subset U$ непустое множество псевдорешений $Z^* \subset W_2^1[a, b]$. Тем самым считаем $D = Z$.

Зададим на Z с помощью функции $g(t)$ — неотрицательной, непрерывной, монотонно возрастающей при $t \geqslant 0$ и подчиненной условию $g(+\infty) = +\infty$ — функционал $\Omega(z) = g(\| z \|_{W_2^1})$. Как отмечалось в § 2.4, функционал Ω тогда обладает необходимыми свойствами, перечисленными в § 1.2. Следствием этого является непустота множества $\bar{Z} \subset W_2^1[a, b]$ нормальных псевдорешений уравнения (2) (см. теорему 1.4.1 и § 2.1). В качестве дополнительного предположения введем требование $\bar{Z} \subset C^1[a, b]$.

Поставим задачу приближенного нахождения нормальных псевдорешений уравнения (2) по конечномерным приближениям для $(A, \bar{u})$. Для этого введем на отрезках $[a, b]$, $[c, d]$ сетки (для простоты равномерные)

$$\omega_N(s) = \{s_j\colon s_j = a + h_s(j-1),\ j = 1, \ldots, N\}, \quad h_s = (b-a)/(N-1),$$
$$\omega_M(x) = \{x_i\colon x_i = c + h_x(i-1),\ i = 1, \ldots, M\}, \quad h_x = (d-c)/(M-1),$$

а в качестве пространств Z_N, U_M будем рассматривать пространства сеточных функций, определенных соответственно на $\omega_N(s)$, $\omega_M(x)$: $\hat{z} = (z_1, \ldots, z_N) \in Z_N$, $\hat{u} = (u_1, \ldots, u_M) \in U_M$. В качестве оператора P_N выберем оператор простого сноса на сетку [47, 177]: $P_N z = (z(s_1), \ldots, z(s_N)) \in Z_N$ для любой функции $z(s) \in W_2^1[a, b] = Z$. Операторы $\bar{P}_N$, $\bar{Q}_M$ будут представлять собой непрерывные операторы восполнения значений сеточных функций путем интерполяции [47]. Тогда свойства 1), 2) операторов P, $\bar{P}$ из § 1 выполнены. Конкретные примеры операторов $\bar{P}_N$, для которых будут также обеспечиваться свойства 3), 4) из § 1, будут рассмотрены ниже.

Пусть вместо правой части $\bar{u}(x)$ уравнения (2) задана сеточная функция $\hat{\bar{u}} \equiv (\bar{u}(x_1), \ldots, \bar{u}(x_M))$ и определен простейший оператор $\bar{Q}_M\colon U_M \to U$ по правилу

$$(\bar{Q}_M \hat{u})(x) = \{\bar{u}(x_i)\colon x_i \leqslant x < x_{i+1},\ i = 1, \ \ldots, \ M-1;\ \bar{u}(x_M)\colon x = x_M\}.$$

Учитывая принадлежность $\bar{u} \in C^1[c, d]$, можно тогда получить конкретное условие аппроксимации вида (4.3) путем следующей выкладки:

$$\rho(\bar{u}, \bar{Q}_M \hat{\bar{u}}) = \| \bar{u} - \bar{Q}_M \hat{\bar{u}} \|_{L_2[c, d]} = \left\{ \sum_{i=1}^{M-1} \int_{x_i}^{x_{i+1}} [\bar{u}(x) - \bar{u}(x_i)]^2 \, dx \right\}^{1/2} \leqslant$$

$$\leqslant \left\{ \sum_{i=1}^{M-1} \int_{x_i}^{x_{i+1}} \| \bar{u} \|_{C^1}^2 (x - x_i)^2 \, dx \right\}^{1/2} \leqslant \| \bar{u} \|_{C^1} \left\{ \sum_{i=1}^{M-1} \frac{h_x^3}{3} \right\}^{1/2} =$$

$$= h_x \| \bar{u} \|_{C^1} \sqrt{\frac{d-c}{3}} \equiv \xi\left(\frac{1}{M}, 0\right) \equiv \xi\left(\frac{1}{M}\right). \tag{3}$$

Функция $\xi(1/M)$ есть мера конечномерной аппроксимации правой части $\bar{u}(x)$ сеточной функции $\hat{\bar{u}}$.

Зададим также конечномерный оператор

$$\hat{A}_H[x_i, \hat{z}] \equiv \hat{A}_H \hat{z} = \sum_{j=1}^{N} K(x_i, s_j, z_j) h_s, \tag{4}$$

определенный при $x_i \in \omega_M(x)$ и для $\hat{z} \in Z_N$ действующий в U_M, непрерывный и представляющий собой простейшую конечномерную аппроксимацию для оператора (1). Получим конкретное условие аппроксимации вида (1.4) для конечномерного оператора (4).

По определению оператора $\bar{Q}_M$ выполнено равенство

$$\rho^2(A[x, z(s)], \bar{Q}_M \hat{A}_H P_N z) =$$

$$= \| A[x, z(s)] - \bar{Q}_M \hat{A}_H P_N z \|^2_{L_2[c, d]} =$$

$$= \sum_{i=1}^{M-1} \int_{x_i}^{x_{i+1}} \left\{ \int_a^b K(x, s, z(s))\, ds - \sum_{j=1}^{N} h_s K(x_i, s_j, z(s_j)) \right\}^2 dx \equiv$$

$$\equiv \sum_{i=1}^{M-1} \int_{x_i}^{x_{i+1}} Q_i^2(x)\, dx. \tag{4}$$

Оценим величину $Q_i(x)$ для $x \in [x_i, x_{i+1}]$ $(i = 1, \ldots, M-1)$:

$$|Q_i(x)| \leqslant \sum_{j=1}^{N-1} \int_{s_j}^{s_{j+1}} |K(x, s, z(s)) - K(x_i, s_j, z(s_j))|\, ds +$$

$$+ |K(x_i, s_N, z(s_N))|\, h_s \equiv Q_{i1} + Q_{i2}. \tag{5}$$

Проще всего оценивается величина Q_{i2} — согласно свойству б) функции K:

$$Q_{i2} \leqslant \| K \|_{C(\Pi_0)} h_s \leqslant \xi_0(\|z\|_C) h_s \leqslant h_s \xi_0(k \|z\|_{W_2^1}). \tag{6}$$

Здесь константа $k > 0$ взята из известной оценки $\|z\|_{C[a, b]} \leqslant k \|z\|_{W_2^1[a, b]}$, справедливой для любой функции $z(s) \in W_2^1[a, b]$ (см. [211, с. 104] и § 0.4).

Для того чтобы оценить Q_{i1}, используем свойства функции K, по которым

$$|K(x, s, z(s)) - K(x_i, s_j, z(s_j))| \leqslant$$

$$\leqslant \| K'_x \|_{C(\Pi_0)} (x - x_i) + \| K'_s \|_{C(\Pi_0)} (s - s_j) + \| K'_z \|_{C(\Pi_0)} |z(s) - z(s_j)| \leqslant$$

$$\leqslant \xi_1(k \|z\|_{W_2^1}) h_x + \xi_2(k \|z\|_{W_2^1})(s - s_j) + \xi_3(k \|z\|_{W_2^1}) |z(s) - z(s_j)|$$

при $x \in [x_i, x_{i+1}]$, $s \in [s_j, s_{j+1}]$. Тогда

$$Q_{i1} \leqslant \xi_1(k \|z\|_{W_2^1}) h_x (b - a) + \xi_2(k \|z\|_{W_2^1}) \sum_{j=1}^{N-1} \int_{s_j}^{s_{j+1}} (s - s_j)\, ds +$$

$$+ \xi_3(k \|z\|_{W_2^1}) \sum_{j=1}^{N-1} \int_{s_j}^{s_{j+1}} |z(s) - z(s_j)|\, ds =$$

$$= (b - a) \left[h_x \xi_1(k \|z\|_{W_2^1}) + \frac{1}{2} \xi_2\left(k \|z\|_{W_2^1}\right) h_s \right] +$$

$$+ \sum_{j=1}^{N-1} \int_{s_j}^{s_{j+1}} |z(s) - z(s_j)|\, ds \cdot \xi_3(k \|z\|_{W_2^1}). \tag{7}$$

В этой выкладке остается оценить последний член. По условию $z(s)\in W_2^1[a, b]$, и, следовательно, функция $z(s)$ абсолютно непрерывна. Тогда, используя свойства абсолютно непрерывных функций, указанные в теореме 0.4.10, и неравенство Коши — Буняковского, получаем оценку

$$\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}|z(s)-z(s_j)|\,ds=\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}\left|\int_{s_j}^{s}z'(t)\,dt\right|ds\leqslant$$

$$\leqslant\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}\left[\int_{s_j}^{s}|z'(t)|\,dt\right]ds\leqslant\sum_{j=1}^{N-1}h_s\int_{s_j}^{s_{j+1}}|z'(t)|\,dt=$$

$$=h_s\int_a^b|z'(t)|\,dt\leqslant h_s\sqrt{b-a}\left\{\int_a^b|z'(t)|^2\,dt\right\}^{1/2}\leqslant h_s\sqrt{b-a}\,\|z\|_{W_2^1}. \tag{8}$$

Поэтому из (7), (8) вытекает неравенство

$$Q_{i1}\leqslant(b-a)\left[\xi_1(k\|z\|_{W_2^1})h_x+\xi_2(k\|z\|_{W_2^1})h_s/2\right]+$$
$$+\sqrt{b-a}\,\|z\|_{W_2^1}\xi_3(k\|z\|_{W_2^1})h_s,$$

учитывая которое, получаем из (4)—(6)

$$\|A[x, z(s)]-\bar{Q}_M\hat{A}_H P_N z\|_{L_2}\leqslant$$
$$\leqslant\sqrt{d-c}\left\{\left[\xi_0(k\|z\|_{W_2^1})+\xi_2(k\|z\|_{W_2^1})(b-a)/2+\right.\right.$$
$$\left.\left.+\sqrt{b-a}\,\|z\|_{W_2^1}\xi_3(k\|z\|_{W_2^1})\right]h_s+(b-a)h_x\xi_1(k\|z\|_{W_2^1})\right\}.$$

Таким образом, можно принять в качестве меры конечномерной аппроксимации из (1.4) в данной задаче функцию

$$\hat{\psi}(H, \Omega)\equiv\sqrt{d-c}\left\{\left[\xi_0[kg^{-1}(\Omega)]+\xi_2[kg^{-1}(\Omega)](b-a)/2+\right.\right.$$
$$\left.\left.+\sqrt{b-a}\,g^{-1}(\Omega)\xi_3[kg^{-1}(\Omega)]\right]h_s+(b-a)\xi_1[kg^{-1}(\Omega)]h_x\right\}, \tag{9}$$
$$H\equiv(h_x,\ h_s).$$

Используя свойства функций g, ξ_m $(m=0, 1, 2, 3)$, введенные выше, нетрудно убедиться, что мера конечномерной аппроксимации $\hat{\psi}$ удовлетворяет стандартным требованиям, предъявленным в § 1.2.

Для проверки возможности использования для решения операторного уравнения (2) алгоритмов, изученных в § 2—4, остается задать оператор $\bar{P}_N$: $Z_N\to Z$. Пусть для определенности это будет оператор кусочно линейной интерполяции значений сеточной функции $\hat{z}\in Z_N$ между узлами:

$$(\bar{P}_N\hat{z})(s)=\{(z_{j+1}-z_j)(s-s_j)/h_s+z_j,\ s\in[s_j, s_{j+1}]\},\quad j=1, \ldots, N-1.$$

Тогда свойство 3) из § 1 выполнено, так как очевидно включение $\bar{P}Pz\in W_2^1[a, b]$ для любой функции $z\in W_2^1[a, b]$. Остается проверить свойство 4) из § 1.

Л е м м а 1. *Для любой функции* $z(s)\in C^1[a, b]$ *справедливо соотношение*

$$\lim_{N\to\infty}\|\bar{P}_N P_N z\|_{W_2^1[a, b]}=\|z\|_{W_2^1[a, b]}.$$

Доказательство. Прежде всего заметим, что для любого элемента $\hat{z}=(z_1, ..., z_N)\in Z_N$ выполнено равенство

$$\|\bar{P}_N\hat{z}\|^2_{W_2^1}=$$
$$=\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}\left\{\frac{(z_{j+1}-z_j)(s-s_j)}{h_s}+z_j\right\}^2 ds+\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}\left(\frac{z_{j+1}-z_j}{h_s}\right)^2 ds=$$
$$=\sum_{j=1}^{N-1} z_j^2 h_s+\sum_{j=1}^{N-1}\left[\frac{(z_{j+1}-z_j)^2}{3}+z_j(z_{j+1}-z_j)\right]h_s+\sum_{j=1}^{N-1}\left(\frac{z_{j+1}-z_j}{h_s}\right)^2 h_s.$$

Из этого равенства, положив $\hat{z}=P_N z$ для $z\in C^1\,[a, b\,]$, получим

$$\|\bar{P}_N P_N z\|^2_{W_2^1}=\sum_{j=1}^{N-1} z^2(s_j)h_s+h_s\sum_{j=1}^{N-1}\left\{\frac{1}{3}[z'(s_j^*)]^2 h_s^2+z(s_j)z'(s_j^*)h_s\right\}+$$
$$+\sum_{j=1}^{N-1}[z'(s_j^*)]^2 h_s\equiv r_1(N)+r_2(N)+r_3(N).$$

Здесь использована формула конечных приращений: $z(s_{j+1})-z(s_j)=$ $=z'(s_j^*)h_s$ $(s_j^*\in[s_j, s_{j+1}\,])$. Остается исследовать величины $r_l(N)$ $(l=1,\ 2,\ 3)$.

Поскольку $z(s)\in C\,[a, b\,]$, то

$$|r_2(N)|\leqslant h_s^2\|z\|_{C^1}(b-a)/3+h_s\|z\|_C\|z\|_{C^1}(b-a),$$

откуда ясно, что $r_2(N)\to 0$ при $N\to\infty$, т. е. при $h_s\to 0$. Далее вследствие интегрируемости функций $z(s)$, $z'(s)$ на отрезке $[a,b\,]$ величины $r_{1,3}(N)$, которые представляют собой интегральные суммы, сходятся при $h_s\to 0$ к своим интегралам:

$$\lim_{N\to\infty} r_1(N)=\int_a^b z^2(s)\,ds,\quad \lim_{N\to\infty} r_3(N)=\int_a^b [z'(s)]^2\,ds.$$

Таким образом, окончательно получим

$$\lim_{N\to\infty}\|\bar{P}_N P_N z\|^2_{W_2^1[a,\,b\,]}=\int_a^b \{z^2(s)+[z'(s)]^2\}\,ds=\|z\|^2_{W_2^1}.$$

Лемма доказана.

Справедливость свойства 4) из § 1 вытекает из предположения о том, что $\bar{Z}\subset C^1\,[a, b\,]$, вида функционала $\Omega(z)$ и леммы 1:

$$\lim_{N\to\infty}\Omega(\bar{P}_N P_N\bar{z})=\lim_{N\to\infty} g(\|\bar{P}_N P_N\bar{z}\|_{W_2^1})=g(\|\bar{z}\|_{W_2^1})=\bar{\Omega}.$$

Резюмировав сказанное о данном примере конечномерной аппроксимации, подчеркнем, что выполнены все условия постановки задачи, сформулированной в § 4. Поэтому для нахождения по данным $\{\hat{A}_H,\ \hat{\bar{u}},\ h_s(N),\ h_x(M),\ \xi,\ \hat{\psi}\}$ приближений к множеству нормальных псевдорешений $\bar{Z}$ уравнения (2) можно применять алгоритмы к.о.п.н., к.о.п.к., к.о.п.с.ф. Отметим, что в этом примере они основаны на минимизации сглаживающего функционала вида

$$\hat{M}^\alpha[z]=\alpha g(\|\bar{P}_N\hat{z}\|_{W_2^1})+f(\|Q_M\hat{A}_H\hat{z}-Q_M\hat{u}\|_{L_2})$$

на множестве Z_N. Согласно общим теоремам сходимости приближений, приведенным в § 2, 3, семейства $\{\bar{P}_N \hat{z}_H\}$, где $H=(h_x, h_s)$, приближенных решений, которые получены по указанным алгоритмам, обладают свойствами $\bar{P}_N \hat{z}_H \xrightarrow{W_2^1} \bar{Z}$, $\|\bar{P}_N \hat{z}_H\|_{W_2^1} \to g^{-1}(\bar{\Omega})$ при $H=(h_x, h_s)\to 0$ (т. е. при M, $N\to\infty$). Как отмечалось в § 2.4, из этих сходимостей вытекает сильная сходимость приближений: $\bar{P}_N \hat{z}_H \overset{W_2^1}{\to} \bar{Z}$ при $H\to 0$.

Отметим, что аналогичным образом можно рассматривать более сложные квадратурные аппроксимации оператора (1) и другие виды операторов $\bar{Q}$, $\bar{P}$. При этом меняются коэффициенты в формулах (3), (9) для мер аппроксимации.

Пример 2. Рассмотрим частный случай уравнения (2):

$$A[x, z(s)] \equiv \int_a^b K_0(x, s) f_0[s, z(s)]\, ds = \bar{u}(x), \quad x \in [c, d], \tag{10}$$

квазирешения которого ищутся на множестве

$$D_+ = \{z(s) \in W_2^1[a, b]: z(s) \geqslant 0\}.$$

При этом предполагается, что $K_0(x, s) \in C^1(\Pi)$, $\Pi \equiv [c, d]\times[a, b]$, $f_0(s, z) \in C(\Pi_1)$, $\Pi_1 \equiv [a, b]\times \mathbf{R}$, и справедливы оценки

$$|f(s, z)| \leqslant \xi(|z|), \quad |f'_s(s, z)| \leqslant \xi_1(|z|), \quad |f'_z(s, z)| \leqslant \xi_2(|z|)$$

для любых $(s, z) \in \Pi_1$. Функции ξ, $\xi_{1,2}$, как и в примере 1, непрерывны и монотонно не убывают.

Если положить $Z \equiv W_2^1[a, b]$, $U \equiv L_2[c, d]$, то, рассуждая, как в примере 1, с учетом результатов примера 2.4.3, можно убедиться, что оператор A из уравнения (10) будет усиленно непрерывным из Z в U (и из D_+ в U).

Предположим, что множество Z^* квазирешений уравнения (10) на множестве D_+ непусто и множество нормальных квазирешений $\bar{Z}$ этого уравнения состоит из непрерывно дифференцируемых на $[a, b]$ функций.

Принимая схему конечномерной аппроксимации из примера 1, зададим конечномерный непрерывный оператор из Z_N в U_M:

$$\hat{A}_H[x_i, \hat{z}] = \hat{A}_H \hat{z} = \sum_{j=1}^{N} K_0(x_i, s_j) f_0[s_j, z_j] h_s,$$
$$x_i \in \omega_M(x).$$

Как и в примере 1, можно получить условие аппроксимации

$$\|A[x, z(s)] - \bar{Q}_M \hat{A}_H P z\|_{L_2[c, d]} \leqslant \hat{\psi}(H, \Omega),$$

$$\begin{aligned}\hat{\psi}(H, \Omega) \equiv \sqrt{d-c}\,\{[(\|K_0\|_{C(\Pi)} + (b-a)\|K'_{0s}\|_{C(\Pi)}/2)\,\xi[kg^{-1}(\Omega)] +\\ + \xi_1[kg^{-1}(\Omega)]\,\|K_0\|_{C(\Pi)}(b-a)/2 +\\ + \sqrt{b-a}\,\|K_0\|_{C(\Pi)}\, g^{-1}(\Omega)\,\xi_2[kg^{-1}(\Omega)]]\, h_s +\\ + (b-a)\|K'_x\|_{C(\Pi)}\,\xi[kg^{-1}(\Omega)]\, h_x\},\end{aligned}$$

в котором мера аппроксимации удовлетворяет стандартным требованиям. Остальные условия применимости алгоритмов конечномерных обобщенных принципов, как и в примере 1, выполнены. Для данного примера эти алгоритмы также дают сильную сходимость приближений к $\bar{Z}$ в $W_2^1[a, b]$.

Пример 3. На множестве $\Pi\equiv\{(x, s): 0\leqslant x\leqslant T,\ 0\leqslant s\leqslant x\}\times\mathbf{R}_+$ задана функция $K(x, s, \gamma)$, удовлетворяющая условиям: а) $K(x, s, \gamma)\in C(\Pi)$; б) $K(x, s, \gamma)$ дифференцируема по x, s при каждом фиксированном $\gamma\geqslant 0$; в) $|K(x, s, \gamma)|\leqslant\xi_0(\gamma)$, $|K'_x(x, s, \gamma)|\leqslant\xi_1(\gamma)$, $|K'_s(x, s, \gamma)|\leqslant\xi_2(\gamma)$ для любых $(x, s, \gamma)\in\Pi$. Входящие в эти оценки функции $\xi_m(\gamma)$ $(m=0, 1, 2)$ непрерывны, а функции $\xi_{1,2}(\gamma)$ еще и монотонно возрастают при $\gamma\geqslant 0$, причем $\xi_{1,2}(+\infty)=+\infty$.

Определим рефлексивное банахово пространство $Z=W_2^1[0, T]\times\mathbf{R}$ с элементами $z\equiv(z(s), \gamma)$ и с нормой $\|z\|^2\equiv\|z(s)\|^2_{W_2^1[0, T]}+\gamma^2$. Зададим также $D\equiv\{z\in Z: z(s)\geqslant 0,\ \gamma\geqslant 0\}$, $U=L_2[0, T]$. Тогда на D определен оператор

$$A[x, z]\equiv A[x; z(s), \gamma]=\int_0^x K(x, s, \gamma)z(s)\,ds,$$

действующий из D в U. Пусть топология τ в Z порождается слабой сходимостью в $W_2^1[0, T]$ и сходимостью в $\mathbf{R}$. Тогда из неравенства

$$|K(x, s, \gamma_1)z_1(s)-K(x, s, \gamma_2)z_2(s)|\leqslant$$
$$\leqslant\|z_1\|_C\,|K(x, s, \gamma_1)-K(x, s, \gamma_2)|+\xi_0(\gamma_2)\|z_1-z_2\|_C,$$

которое справедливо для любых $z_{1,2}\in Z$, вытекает с учетом условий на функции K, ξ_0 непрерывность оператора A из $C[0, T]\times\mathbf{R}$ в U и усиленная непрерывность этого оператора из Z в U.

Введем теперь уравнение

$$A[x; z(s), \gamma]=\bar{u}(x),\quad x\in[0, T],\tag{11}$$

относительно неизвестной функции $z(s)\in W_2^1[0, T]$, где $z(s)\geqslant 0$, и неизвестного параметра $\gamma\geqslant 0$. Предположим, что правой части $\bar{u}(x)\in C^1[0, T]$ уравнения (11) отвечает решение $\bar{z}\equiv(\bar{z}(s), \bar{\gamma})$ этого уравнения, причем $\bar{z}(s)\in C^1[0, T]$, $\bar{z}(s)\geqslant 0$, $\bar{\gamma}\geqslant 0$, т. е. $\bar{z}\in D$.

Для поиска приближений к $\bar{z}$ зададим функционал $\Omega(z)\equiv\Omega[z(s), \gamma]=\|z\|^2_{W_2^1[0, T]}+p_0\xi_0^2(\gamma)+p_1\xi_1^2(\gamma)+p_2\xi_2^2(\gamma)$, где константы p_0, p_1, $p_2>0$ будут указаны ниже. Функционал $\Omega(z)$ обладает свойствами, перечисленными в § 1.2 и используемыми в § 4. Действительно, норма $\|z(s)\|_{W_2^1}$ слабо полунепрерывна снизу в гильбертовом пространстве $W_2^1[0, T]$ (см., например, § 0.3), а функции $\xi_{0,1,2}(\gamma)$ непрерывны при $\gamma\geqslant 0$. Это доказывает τ-секвенциальную полунепрерывность снизу функционала Ω на τ-замкнутом в Z множестве D. Далее из неравенства $\Omega(z)\leqslant C^2$ следует, что $\|z(s)\|_{W_2^1[0, T]}\leqslant C^2$, $0\leqslant\gamma\leqslant\max\{\xi_1^{-1}(C/\sqrt{p_1}),\ \xi_2^{-1}(C/\sqrt{p_2})\}$. Поэтому вследствие слабой компактности шара в $W_2^1[0, T]$ и компактности в $\mathbf{R}$ отрезка множество $\Omega_{C^2}\equiv\{z\in D: \Omega(z)\leqslant C^2\}$ τ-секвенциально компактно.

Обратимся теперь к конечномерной аппроксимации уравнения (11). Пусть задана сетка

$$\omega_N(s)=\omega_M(x)=\{s_j\colon s_j=H(j-1),\, j=1, \ldots, N\} \quad H=\frac{T}{N-1}, \quad M=N,$$

и на ней задано пространство сеточных функций $Z'_N \equiv U_N$: $\hat{z}= =(z_1, \ldots, z_N)\in Z'_N$. Введем пространство $Z_N \equiv Z'_N \times \mathbf{R}$ с элементами $(\hat{z}, \gamma)$ и определим на Z оператор P_N по правилу

$$P_N z=(z(s_1), \ldots, z(s_N), \gamma) \quad \forall z=(z(s), \gamma)\in Z,\ s_1, \ldots, s_N \in \omega_N(s).$$

Оператор $\bar{P}_N$, действующий из Z_N в Z, будет ставить в соответствие элементу $(\hat{z}, \gamma)\in Z_N$ пару $(\hat{z}(s), \gamma)\in Z$, в которой функция $\hat{z}(s)\in \in W_2^1\,[0, T\,]$ получается путем линейной интерполяции значений сеточной функции $\hat{z}\in Z'_N$ между узлами. Оператор $\bar{Q}_N\colon U_N \to U$ определим по правилу

$$(\bar{Q}_N \hat{u})(x)=\{u_1\colon x\in[x_1, x_2]; \quad u_i\colon x\in(x_i, x_{i+1}],\ i=2, \ldots, N-1\}.$$

Зададим конечномерный оператор

$$\hat{A}_H(\hat{z}, \gamma)=\sum_{j=1}^{i} K(x_i, s_j, \gamma)\, z_j H,$$

непрерывный из Z_N в U_N. Используя условия на функцию K, а также принадлежность $z(s)\in W_2^1\,[0, T\,]$, можно по схеме, которая применена в примере 1, получить оценку

$$\|A[x; z(s), \gamma]-\bar{Q}_N \hat{A}_H P_N z\|_{L_2[0, T]} \leqslant \hat{\psi}(H, \Omega), \tag{12}$$

$$\hat{\psi}(H, \Omega)\equiv\sqrt{T}h\,\{\|z(s)\|^2_{W_2^1[0, T]}+p_0\xi_0^2(\gamma)+p_1\xi_1^2(\gamma)+p_2\xi_2^2(\gamma)\}/2= =\sqrt{T}h\,\Omega(z)/2.$$

При этом принято $p_0=3(\sqrt{T}+k)^2$, $p_1=p_2=3k^2T^2$. Мера аппроксимации $\hat{\psi}(H, \Omega)$ удовлетворяет предположениям из § 1.2.

Аппроксимация правой части уравнения (11) проводится, как в примере 1.

Сказанное позволяет утверждать, что для поиска решения $\bar{z}=(\bar{z}(s), \bar{\gamma})$ уравнения (11) можно применять модифицированные конечномерные алгоритмы из § 4. При этом с учетом сказанного в примерах 1, 2 получается сходимость приближений в норме пространства Z к $\bar{z}$ при $H\to 0$.

В качестве иллюстрации рассмотрим уравнение (11) с ядром $K(x, s, \gamma)=\exp\{-\gamma(x-s)\}$ $(\gamma\geqslant 0)$. Тогда, как легко видеть, все условия на функцию K справедливы, причем $\xi_0(\gamma)=1$, $\xi_{1,2}(\gamma)=\gamma$. Можно показать, что для такого уравнения можно в качестве функционала Ω взять квадратичный по $z(s)$, γ функционал

$$\Omega(z)=\|z(s)\|^2_{W_2^1[0, T]}+6k^2T^2\gamma^2.$$

Тогда мера аппроксимации из (12) примет вид

$$\hat{\psi}(H, \Omega)=\sqrt{T}h\,\{\Omega+p_0\}/2.$$

Можно привести и другие примеры задач, для решения которых применимы конечномерные обобщенные принципы невязки, квазирешений и сглаживающего функционала, а также модификации рассмотренных примеров с другими видами аппроксимации, сходимости и т. д.

В заключение главы сделаем некоторые замечания по поводу численной реализации алгоритмов, данных в § 2—4. Методика этой реализации фактически изложена в § 1.14. Отличие заключается лишь в использовании конечномерных функционалов. На основе этой

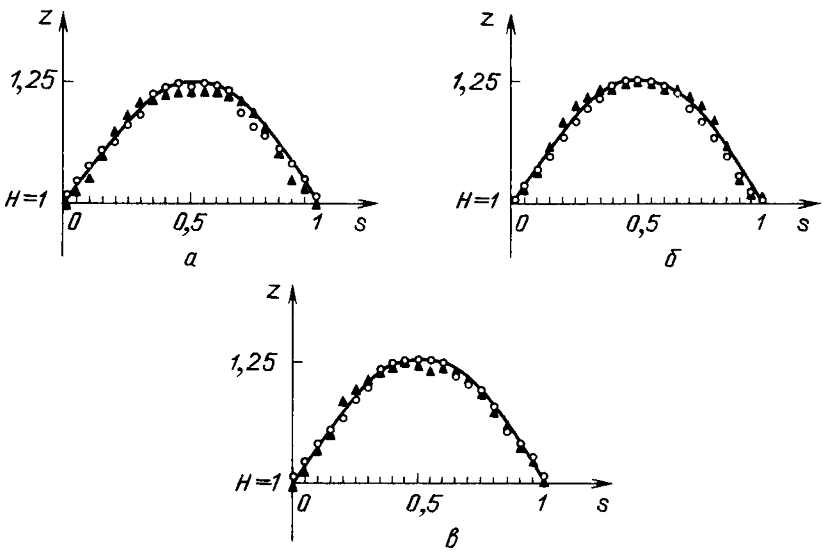

Рис. 16. Применение алгоритмов обобщенных принципов для решения интегрального уравнения (13): *а* — к.о.п.н.; *б* — к.о.п.к.; *в* — к.о.п.с.ф. Точное решение: $\bar{z}$(—); приближенное решение: $\sigma=10^{-4}$ (○○○); $\sigma=10^{-3}$ (▲▲▲)

методики был создан комплекс программ для решения нелинейных интегральных уравнений вида (2), (10), (11) и некоторых других [135].

При решении задач такого рода в качестве базового метода минимизации сглаживающего функционала, а также метода приближенного определения обобщенной меры несовместности использовалась комбинация метода Флетчера — Ривса [170] (с некоторыми модификациями) и метода Ньютона. Такой комбинированный метод хорошо зарекомендовал себя с точки зрения скорости сходимости.

Приведем примеры решения с помощью упомянутого комплекса программ некоторых модельных задач. В теории нелинейных некорректных задач в качестве стандартного примера обычно используется интегральное уравнение типа (2), возникающее при решении некоторых вопросов геофизики [205]:

$$\int_a^b \ln\frac{(x-\xi)^2+H^2}{(x-\xi)^2+(\bar{z}(\xi)-H)^2}\,d\xi=\bar{u}(x), \quad x\in[c,\ d]. \qquad (13)$$

В этом уравнении параметр $H>0$ считается заданным, а искомое решение есть $\bar{z}(s)$. При дополнительных предположениях о том, что $\bar{z}(\xi)\in W_2^1[a, b]$, $\bar{z}(\xi)\geqslant H+\varepsilon$, где $\varepsilon>0$ — заданное число, и после замены $z^2(\xi)=\bar{z}(\xi)-H-\varepsilon$ уравнение для $z(\xi)$, которое получалось из (13) аппроксимировалось по схеме примера 1 с $M=N=41$, $a=0$, $b=1$, $c=-1$, $d=2$. В качестве модельного решения выбиралась функция $\bar{z}(\xi)=\xi(1-\xi)+H+\varepsilon$, определенная при $0\leqslant\xi\leqslant 1$. По ней вычислялась функция $\bar{u}(x)$ при $x\in[-1, 2]$, которая затем возмущалась помехой вида $\sigma\omega$, где ω — равномерно распределенная на $[-1, 1]$ случайная величина. Функционал Ω задавался в виде $\Omega(z)=\|z\|^2_{W_2^1[a, b]}$. В алгоритмах § 4 полагалось $f(x)=x^2$, $C=1{,}01$, $q=1{,}01$, $\omega_\eta=0{,}4$, $p=9/10$, $\varepsilon=10^{-5}$. За начальное приближение для метода минимизации принималась постоянная функция $\bar{z}_0(\xi)=1+H+\varepsilon$.

На рис. 16 приведены результаты решения уравнения (13) с помощью алгоритмов к.о.п.н., к.о.п.к., к.о.п.с.ф. соответственно для различных уровней возмущения σ правой части $\bar{u}$. Расчеты проводились на ЭВМ БЭСМ-6 (язык Фортран). Время расчета приближенного решения для представленных вариантов составляет 3—7 мин в зависимости от выбора начальной оценки для α и числа $n(\delta)$ (см. § 1.14).

ГЛАВА 4

КУСОЧНО РАВНОМЕРНАЯ РЕГУЛЯРИЗАЦИЯ НЕКОРРЕКТНЫХ ЗАДАЧ С РАЗРЫВНЫМИ РЕШЕНИЯМИ

При решении некорректных задач методом регуляризации часто возникает проблема построения кусочно равномерных приближений к точному решению, имеющему, например, конечное число точек разрыва первого рода с неизвестным положением. В данной главе рассматривается одно из возможных решений этой проблемы на основе использования функций с ограниченной вариацией [121, 122, 132].

§ 1. Некоторые свойства функций с ограниченной вариацией

Обозначим символом $V[a, b]$ пространство функций, определенных на отрезке $[a, b]$ и имеющих там ограниченную вариацию. Если наделить пространство $V[a, b]$ нормой

$$\|z\| \equiv \|z\|_{V[a,b]} = |z(a)| + \bigvee_a^b(z) \quad \forall z \in V[a, b],$$

то, как известно, $V[a, b]$ становится банаховым пространством. Особенностью функций с ограниченной вариацией является их ограниченность и наличие не более чем счетного множества точек разрыва первого рода (см. § 0.4).

Введем некоторые обозначения. Пусть $s_1, \ldots, s_N$ — точки отрезка $[a, b]$, подчиненные требованию $a = s_1 < \ldots < s_N = b$. Будем обозначать совокупность таких точек символом $\Pi_N = \{s_1, \ldots, s_N\}$ и называть Π_N *разбиением* отрезка $[a, b]$. Пусть $z(s)$ — ограниченная на $[a, b]$ функция, а Π_N — некоторое разбиение. Обозначим

$$\omega(z, \Pi_N) = \sum_{k=1}^{N-1} |z(s_{k+1}) - z(s_k)|.$$

Отметим, что для функции $z(s) \in V[a, b]$ оказывается, что

$$\omega(z, \Pi_N) \leqslant \bigvee_a^b(z), \tag{1}$$

каково бы ни было разбиение Π_N [93, 168].

В дальнейшем нам понадобится ряд утверждений, связанных со сходимостью последовательностей функций ограниченной вариации.

Лемма 1. *Предположим, что последовательность функций* $\{z_n(s)\}_{n=1}^{\infty} \subset V[a, b]$ *обладает следующими свойствами:* а) $z_n(s)$ *сходится при* $n \to \infty$ *для каждого* $s \in [a, b]$ *к некоторой функции* $\bar{z}(s)$; б) $\|z_n\| \leqslant \bar{M} = \text{const}$ *для всякого* n. *Тогда* $\bar{z}(s) \in V[a, b]$, *причем*

$$\overset{b}{\underset{a}{\vee}}(\bar{z}) \leqslant \varliminf_{n \to \infty} \overset{b}{\underset{a}{\vee}}(z_n), \quad \|\bar{z}\| \leqslant \varliminf_{n \to \infty} \|z_n\|.$$

Доказательство. Рассмотрим произвольное разбиение Π_N отрезка $[a, b]$. Тогда из сходимости $z_n(s) \to \bar{z}(s)$ в точках разбиения и из неравенства (1), записанного для z_n, получим

$$\omega(\bar{z}, \Pi_N) = \lim_{n \to \infty} \omega(z_n, \Pi_N) = \varliminf_{n \to \infty} \omega(z_n, \Pi_N) \leqslant \varliminf_{n \to \infty} \overset{b}{\underset{a}{\vee}}(z_n) \leqslant \bar{M}.$$

Поэтому

$$\overset{b}{\underset{a}{\vee}}(\bar{z}) = \sup\{\omega(\bar{z}, \Pi_N) \colon \Pi_N \in \{\Pi_N\}\} \leqslant \varliminf_{n \to \infty} \overset{b}{\underset{a}{\vee}}(z_n) \leqslant \bar{M},$$

что и требуется доказать.

Определение 1. Последовательность функций $\{z_n(s)\} \subset V[a, b]$ называется П-*сходящейся* к элементу $z_0(s)$, если $z_n(s) \to z_0(s)$ при $n \to \infty$ для любого $s \in [a, b]$ и $\|z_n\| \leqslant \bar{M} = \text{const}$ для всякого n.

Ясно, что по лемме 1 $z_0(s) \in V[a, b]$.

Лемма 2. *Пусть последовательность функций* $\{z_n(s)\} \subset V[a, b]$ П-*сходится при* $n \to \infty$ *к функции* $\bar{z}(s) \in V[a, b]$ *и, кроме того,* $\|z_n\| \leqslant \|\bar{z}\|$ *для любых* n. *Тогда* $\|z_n\| \to \|\bar{z}\|$.

Доказательство. Из условий леммы 2 и по лемме 1 легко получить неравенства

$$\|\bar{z}\| \leqslant \varliminf_{n \to \infty} \|z_n\| \leqslant \varlimsup_{n \to \infty} \|z_n\| \leqslant \|\bar{z}\|,$$

откуда и следует доказываемое предельное соотношение.

Лемма 3. *Если для последовательности* $\{z_n(s)\}$ *выполнены условия леммы* 2, *то, каковы бы ни были числа* s_1, s_2: $a \leqslant s_1 < s_2 \leqslant b$, *справедливо соотношение*

$$\lim_{n \to \infty} \overset{s_2}{\underset{s_1}{\vee}}(z_n) = \overset{s_2}{\underset{s_1}{\vee}}(\bar{z}).$$

Доказательство. Поскольку выполнены условия леммы 2, а значит, и леммы 1, то можно утверждать, что для любых чисел s'_1, s'_2: $a \leqslant s'_1 < s'_2 \leqslant b$ справедливы неравенства

$$\varlimsup_{n \to \infty} \overset{s'_2}{\underset{s'_1}{\vee}}(z_n) \geqslant \varliminf_{n \to \infty} \overset{s'_2}{\underset{s'_1}{\vee}}(z_n) \geqslant \overset{s'_2}{\underset{s'_1}{\vee}}(\bar{z}). \tag{2}$$

Предположим теперь, что для $s'_1 = s_1$, $s'_2 = s_2$

$$\varlimsup_{n \to \infty} \overset{s_2}{\underset{s_1}{\vee}}(z_n) > \overset{s_2}{\underset{s_1}{\vee}}(\bar{z}). \tag{3}$$

Для определенности считаем s_1, s_2 внутренними точками отрезка $[a, b]$. Найдется такая последовательность функций $\{z_{n_l}\} \subset \{z_n\}$, что

$$\overline{\lim_{n \to \infty}} \overset{s_2}{\underset{s_1}{\vee}}(z_n) = \lim_{l \to \infty} \overset{s_2}{\underset{s_1}{\vee}}(z_{n_l}) > \overset{s_2}{\underset{s_1}{\vee}}(\bar{z}). \qquad (4)$$

Из условия ограниченности норм $\|z_n\| \leqslant \|\bar{z}\|$ вытекает ограниченность числовых последовательностей

$$\left\{ \overset{s_1}{\underset{a}{\vee}}(z_{n_l}) \right\}, \quad \left\{ \overset{b}{\underset{s_2}{\vee}}(z_{n_l}) \right\},$$

и потому можно выбрать такую подпоследовательность $\{z_{n_{l_r}}(s)\} \subset \{z_{n_l}(s)\}$, для которой существуют пределы

$$\lim_{r \to \infty} \overset{s_1}{\underset{a}{\vee}}(z_{n_{l_r}}) \geqslant \overset{s_1}{\underset{a}{\vee}}(\bar{z}), \quad \lim_{r \to \infty} \overset{b}{\underset{s_2}{\vee}}(z_{n_{l_r}}) \geqslant \overset{b}{\underset{s_2}{\vee}}(\bar{z}). \qquad (5)$$

Оценки для этих пределов следуют из (2). Далее,

$$\lim_{r \to \infty} \overset{b}{\underset{a}{\vee}}(z_{n_{l_r}}) = \lim_{r \to \infty} \left[\overset{s_1}{\underset{a}{\vee}}(z_{n_{l_r}}) + \overset{s_2}{\underset{s_1}{\vee}}(z_{n_{l_r}}) + \overset{b}{\underset{s_2}{\vee}}(z_{n_{l_r}}) \right] > \overset{s_1}{\underset{a}{\vee}}(\bar{z}) + \overset{s_2}{\underset{s_1}{\vee}}(\bar{z}) + \overset{b}{\underset{s_2}{\vee}}(\bar{z}) = \overset{b}{\underset{a}{\vee}}(\bar{z}),$$

где использованы (4), (5). Последнее соотношение означает, что

$$\lim_{r \to \infty} \|z_{n_{l_r}}\| > \|\bar{z}\|,$$

а это противоречит выводу леммы 2:

$$\lim_{n \to \infty} \|z_n\| = \lim_{r \to \infty} \|z_{n_{l_r}}\| = \|\bar{z}\|.$$

Итак, предположение (3) о строгости неравенства приводит к противоречию. Значит,

$$\overline{\lim_{n \to \infty}} \overset{s_2}{\underset{s_1}{\vee}}(z_n) = \overset{s_2}{\underset{s_1}{\vee}}(\bar{z}),$$

что вместе с (2) для $s'_1 = s_1$, $s'_2 = s_2$ и доказывает лемму.

Следствие 1. *Если в качестве условий леммы 3 принять требование П-сходимости последовательности $\{z_n\}$ к $\bar{z}$, а также условие*

$$\overline{\lim_{n \to \infty}} \|z_n\| \leqslant \|\bar{z}\|,$$

то лемма 3 остается справедливой.

Сформулируем теперь важную теорему, которая будет существенно использоваться при установлении сходимости алгоритмов кусочно-равномерной регуляризации.

Теорема 1. *Пусть задана функциональная последовательность $\{z_n(s)\} \subset V[a, b]$, для которой выполнены условия: 1) она П-сходится при $n \to \infty$ к функции $\bar{z}(s) \subset V[a, b]$; 2) $\|z_n\| \leqslant \|\bar{z}\|$ для каждого n. Тогда $z_n(s)$ сходится к $\bar{z}(s)$ равномерно на любом отрезке $[s', s''] \subset [a, b]$, не содержащем точек разрыва функции $\bar{z}(s)$.*

Доказательство. Допустим, что теорема 1 неверна: пусть на некотором отрезке $[s', s'']\subset[a, b]$, не содержащем точек разрыва функции $\bar{z}(s)$, сходимость не является равномерной. Это означает, что для определенного $\varepsilon_0>0$ найдется последовательность точек $s_n\in[s', s'']$ такая, что $|z_n(s_n)-\bar{z}(s_n)|>\varepsilon_0$, как только $n\geqslant n_0$. Выделим из $\{s_n\}$ сходящуюся подпоследовательность $\{s_{n_k}\}$: $s_{n_k}\to s^*\in[s', s'']$ при $k\to\infty$. Тогда можно написать неравенство

$$\varepsilon_0<|z_{n_k}(s_{n_k})-\bar{z}(s_{n_k})|\leqslant|z_{n_k}(s_{n_k})-z_{n_k}(s^*)|+$$
$$+|z_{n_k}(s^*)-\bar{z}(s^*)|+|\bar{z}(s^*)-\bar{z}(s_{n_k})|\equiv J_{1k}+J_{2k}+J_{3k}. \tag{6}$$

Изучим поведение последовательности чисел $\{J_{jk}\}_{k=1}^{\infty}$ $(j=1, 2, 3)$ при $k\to\infty$. Непрерывность функции $\bar{z}(s)$ на $[s', s'']$ и, в частности, в точке s^* порождает сходимость $J_{3k}\to0$. Из условия 1) теоремы следует также, что $J_{2k}\to0$. Рассмотрим более подробно J_{1k}. Предположим для определенности, что s^* — внутренняя точка отрезка $[s', s'']$. Если это не так, то доказательство незначительно изменяется, оставаясь по идее тем же. Поскольку функция $\bar{z}(s)\in V[a, b]$ непрерывна в точке $s^*\in(s', s'')$, то, согласно известному свойству функций ограниченной вариации (см., например, теорему 0.4.14), в этой же точке непрерывна и функция

$$\bar{q}(s)\equiv\bigvee_a^s(\bar{z}).$$

Таким образом, можно выбрать столь малое число $\delta_0>0$, что $[s^*-\delta_0, s^*+\delta_0]\subset[s', s'']$ и

$$\bar{q}(s^*+\delta_0)-\bar{q}(s^*-\delta_0)=\bigvee_{s^*-\delta_0}^{s^*+\delta_0}(\bar{z})<\frac{\varepsilon_0}{2}. \tag{7}$$

Выберем далее такой номер $N(\delta_0)$, что $|s^*-s_{n_k}|<\delta_0$ при $k>N(\delta_0)$. Тогда, учитывая (1), получаем

$$J_{1k}=|z_{n_k}(s_{n_k})-z_{n_k}(s^*)|\leqslant|\bigvee_{s_{n_k}}^{s^*}(z_{n_k})|\leqslant\bigvee_{s^*-\delta_0}^{s^*+\delta_0}(z_{n_k}),\quad k>N(\delta_0). \tag{8}$$

Условия доказываемой теоремы совпадают с условиями лемм 2, 3. Поэтому для последовательности $\{z_{n_k}\}$ по лемме 3

$$\lim_{k\to\infty}\bigvee_{s^*-\delta_0}^{s^*+\delta_0}(z_{n_k})=\bigvee_{s^*-\delta_0}^{s^*+\delta_0}(\bar{z}). \tag{9}$$

Тем самым, переходя к пределу в (8) и учитывая (7), (9), можно убедиться, что

$$\lim_{k\to\infty}J_{1k}<\varepsilon_0/2.$$

Суммируя результаты о сходимости J_{jk} $(j=1, 2, 3)$ в формуле (6), приходим к противоречивому неравенству

$$\varepsilon_0\leqslant\lim_{k\to\infty}(J_{1k}+J_{2k}+J_{3k})<\varepsilon_0/2,$$

что и доказывает теорему.

С л е д с т в и е 2. *Пусть выполнено условие* 1) *теоремы* 1, *а вместо условия* 2) *справедливо соотношение*

$$\overline{\lim_{n\to\infty}}\|z_n\|\leqslant\|\bar{z}\|.$$

Тогда утверждение теоремы 1 *остается справедливым.*

Действительно, из неравенства, доказанного в лемме 1, и из условия следствия 2 непосредственно получаем, что $\|z_n\|\to\|\bar{z}\|$. Тогда по следствию 1 верна лемма 3, которая фактически и использовалась при доказательстве теоремы 1.

В заключение остановимся на одном свойстве функций с ограниченной вариацией.

О п р е д е л е н и е 2. Условимся называть точку разрыва первого рода $s_0\in(a, b)$ функции $z(s)\in V[a, b]$ *обычной*, если выполнено условие

$$\underline{\lim_{s\to s_0}}\, z(s)=\min\left[z(s_0-0),\ z(s_0+0)\right]\leqslant z(s_0)\leqslant\overline{\lim_{s\to s_0}}\, z(s)=$$
$$=\max\left[z(s_0-0),\ z(s_0+0)\right].$$

Если a или b представляют собой точки разрыва функции $z(s)$, то они по определению считаются обычными.

Очевидно, что точка разрыва $s_0\in(a, b)$ является обычной тогда и только тогда, когда выполнено равенство

$$|z(s_0-0)-z(s_0)|+|z(s_0+0)-z(s_0)|=|z(s_0-0)-z(s_0+0)|.$$

Л е м м а 4. *Предположим, что функции* $z_1(s)$, $z_2(s)\in V[a, b]$ *совпадают всюду, кроме, быть может, их общих точек разрыва, лежащих на* (a, b), *которые для функции* $z_1(s)$ *являются обычными. Тогда* $\|z_1\|=\|z_2\|$, *если все точки разрыва функции* $z_2(s)$ *обычные, и* $\|z_1\|<\|z_2\|$, *если это требование нарушается.*

Д о к а з а т е л ь с т в о. Достаточно рассмотреть случай, когда $z_1(s)$, $z_2(s)$ отличаются лишь в одной точке разрыва s_0 $(a<s_0<b)$. Воспользуемся равенством

$$\bigvee_a^b(z)=\lim_{s\to s_0-0}\bigvee_a^s(z)+\lim_{s\to s_0+0}\bigvee_s^b(z)+|z(s_0-0)-z(s_0)|+|z(s_0+0)-z(s_0)|,$$

справедливым для произвольной функции $z\in V[a, b]$. Из него при $z=z_1$ с учетом того, что для обычной точки разрыва выполнено соотношение

$$|z_1(s_0-0)-z_1(s_0)|+|z_1(s_0+0)-z_1(s_0)|=|z_1(s_0-0)-z_1(s_0+0)|,$$

получим

$$\bigvee_a^b(z_1)=\lim_{s\to s_0-0}\bigvee_a^s(z_1)+\lim_{s\to s_0+0}\bigvee_s^b(z_1)+|z_1(s_0-0)-z_1(s_0+0)|.$$

Для функции $z_2(s)$, совпадающей с $z_1(s)$ всюду, кроме точки s_0, аналогично имеем

$$\bigvee_a^b(z_2)=\lim_{s\to s_0-0}\bigvee_a^s(z_2)+\lim_{s\to s_0+0}\bigvee_s^b(z_2)+$$
$$+|z_2(s_0-0)-z_2(s_0)|+|z_2(s_0+0)-z_2(s_0)|=$$
$$=\lim_{s\to s_0-0}\bigvee_a^s(z_1)+\lim_{s\to s_0+0}\bigvee_s^b(z_1)+|z_2(s_0-0)-z_2(s_0)|+|z_2(s_0+0)-z_2(s_0)|.$$

При этом выполнено неравенство

$$|z_2(s_0-0)-z_2(s_0)|+|z_2(s_0+0)-z_2(s_0)|\geqslant$$
$$\geqslant|z_2(s_0-0)-z_2(s_0+0)|=|z_1(s_0-0)-z_1(s_0+0)|,$$

которое будет строгим, если s_0 не является обычной точкой разрыва функции $z_2(s)$, и которое переходит в равенство, если s_0 — обычная точка разрыва этой функции.

Сравнивая теперь вариации функций z_1, z_2, можно убедиться в справедливости заключения леммы 4.

Замечание. Утверждения типа теоремы 1 и лемм 1—3 рассматривались в несколько более общем случае в работе [122] (ср. также [67, 206]).

§ 2. Постановка задачи и алгоритмы ее решения

Пусть A — оператор, действующий из $V[a, b]$ в нормированное пространство U, а F — функционал, определенный на $V[a, b]$.

Определение 1. Оператор A (функционал F) *П-непрерывен на элементе* $z_0\in V[a, b]$, если для любой последовательности $\{z_n(s)\}$, П-сходящейся к $z_0(s)$ при $n\to\infty$, справедливо предельное соотношение $Az_n\xrightarrow{U}Az_0$ $[F(z_n)\to F(z_0)]$. Если A (или F) П-непрерывен для всякого $z_0\in V[a, b]$, то он называется П-непрерывным на $V[a, b]$.

Обозначим $Z\equiv V[a, b]$ и введем в Z топологию τ, порождающую П-сходимость. Пусть далее A — оператор, который П-непрерывен на Z, а $\Omega_1(z)$ — неотрицательный функционал, также П-непрерывный на Z. Зададим, кроме того, функционал $\Omega(z)\equiv\|z\|+p\Omega_1(z)$, где $p\geqslant 0$ — некоторая константа.

Лемма 1. *Функционал $J(z)\equiv\|Az-u\|_U$ τ-непрерывен на Z. Функционал $\Omega(z)$ τ-полунепрерывен снизу на Z.*

Доказательство. П-непрерывность функционала J вытекает из П-непрерывности оператора A. Пусть теперь $\{z_n(s)\}\subset V[a, b]$ — любая последовательность, которая П-сходится к $z_0(s)\in V[a, b]$. По лемме 1.1 с учетом П-непрерывности функционала Ω_1 получается соотношение

$$\varliminf_{n\to\infty}\Omega(z_n)=\varliminf_{n\to\infty}\{\|z_n\|+p\Omega_1(z_n)\}=\varliminf_{n\to\infty}\|z_n\|+p\lim_{n\to\infty}\Omega_1(z_n)\geqslant$$
$$\geqslant\|z_0\|+p\Omega_1(z_0)=\Omega(z_0),$$

которое и означает полунепрерывность снизу на Z функционала Ω относительно П-сходимости.

Лемма 2. *Множество* $\Omega_C \equiv \{z \in V[a, b]: \Omega(z) \leqslant C\}$, *если оно непусто, является* τ*-секвенциально компактным.*

Действительно, из неравенства $\Omega(z_n) \leqslant C$, выполненного для произвольной последовательности $\{z_n\} \subset \Omega_C$, следует равномерная ограниченность норм $\|z_n\| \leqslant C$. Тогда по теореме Хелли о компактности (см. теорему 0.4.15) последовательность $\{z_n\}$ компактна относительно поточечной сходимости, и поэтому τ-секвенциально компактна.

Рассмотрим операторное уравнение

$$Az = u, \quad z \in V[a, b], \quad u \in U, \tag{1}$$

и будем полагать множество Z^* его псевдорешений, отвечающих правой части $\bar{u} \in U$, непустым.

Из общей теоремы 1.4.1 на основании результатов лемм 1, 2 можно сделать вывод, что множество $\bar{Z}$ Ω-оптимальных псевдорешений уравнения (1) на классе $V[a, b]$ непусто.

В дальнейшем будем считать, что выполнено *обобщенное свойство единственности* Ω-оптимального псевдорешения: все Ω-оптимальных псевдорешения уравнения (1) совпадают между собой поточечно на $[a, b]$, за исключением, быть может, их общих точек разрыва, которые, если они имеются, предполагаются обычными. Это предположение ведет по лемме 1.5 к совпадению норм всех Ω-оптимальных псевдорешений, а в силу равенств

$$\Omega(z) = \|z\| + p\Omega_1(z), \quad \Omega(z) = \bar{\Omega} \quad \forall z \in \bar{Z}$$

и к постоянству значения функционала Ω_1 на множестве $\bar{Z}$. Для удобства формулировок выделим одно из Ω-оптимальных решений и обозначим его как $\bar{z}(s)$.

Пусть теперь вместо точных данных $(A, \bar{u})$ задачи (1) рассматриваются их приближения (A_h, u_σ), подчиненные требованиям из § 2.1. В данном случае это означает, что оператор A_h, определенный на $V[a, b]$, П-непрерывен и удовлетворяет условию аппроксимации

$$\|A_h z - Az\|_U \leqslant \psi(h, \|z\|) \quad \forall z \in V[a, b],$$

причем число $h \geqslant 0$ и функция ψ известны, а последняя обладает стандартными свойствами, сформулированными в § 1.2. Известные число $\sigma \geqslant 0$ и элемент $u_\sigma \in U$ подчинены требованию $\|\bar{u} - u_\sigma\|_U \leqslant \sigma$.

Задача о кусочно равномерной регуляризации уравнения (1) заключается в построении по данным $(A_h, u_\sigma, h, \sigma, \psi)$ приближений $z_\delta \in V[a, b]$ $(\delta \equiv (h, \sigma))$ к Ω-оптимальным псевдорешениям из множества $\bar{Z}$ так, чтобы для любой последовательности $\{\delta_n\}$, сходящейся к нулю при $n \to \infty$, соответствующая последовательность $\{z_n(s)\}$, где $z_n \equiv z_{\delta_n}$, сходилась к $\bar{z}(s)$ равномерно на любом отрезке $[s', s''] \subset [a, b]$, который не содержит точек разрыва функции $\bar{z}(s)$.

Поскольку для рассмотренной постановки задачи выполнены стандартные требования из § 2.1, то для поиска Ω-оптимальных псевдорешений уравнения (1) можно использовать алгоритмы о.п.н., о.п.к., о.п.с.ф., данные в § 2.2. Они основаны в нашем случае на минимизации сглаживающего функционала

$$M^{\alpha}[z]=\alpha g\left[|z(a)|+\bigvee_{a}^{b}(z)+p\Omega_1(z)\right]+f(\|A_h z-u_\sigma\|),\quad \alpha>0,\quad z\in Z,$$

в котором функция $g(x)$ непрерывна и монотонно возрастает при $x\geqslant 0$ так, что $g(+\infty)=+\infty$, а $f(x)\in\mathscr{F}^m[0,+\infty)$. Пусть $z^{\alpha_\delta}\in V[a,b]$ — приближение к $\bar{Z}$, полученное с помощью указанных алгоритмов. Тогда, согласно общей теории, изложенной в гл. 1, 2 для любой последовательности $\{\delta_n\}$, сходящейся при $n\to\infty$ к нулю, соответствующая последовательность приближенных решений $\{z_n\}$, где $z_n=z^{\alpha}\delta_n$, П-сходится при $n\to\infty$ к $\bar{Z}$ и $\Omega(z_n)\to\bar{\Omega}$. Оказывается, что кроме этого приближения z^{α_δ} дают решение задачи о кусочно равномерной регуляризации. Сформулируем соответствующий результат.

Теорема 1. *Последовательность $\{z_n(s)\}$ сходится к $\bar{z}(s)$ при $n\to\infty$ равномерно на любом отрезке $[s',s'']\subset[a,b]$, не содержащем точек разрыва функции $\bar{z}(s)$.*

Доказательство. Поскольку последовательность $\{z_n\}$ П-сходится к $\bar{Z}$, а элементы множества $\bar{Z}$ имеют одинаковую норму, то по лемме 1.1 с учетом П-непрерывности функционала Ω_1 и сходимости $\Omega(z_n)\to\bar{\Omega}$ при $n\to\infty$ можно получить

$$\|\bar{z}\|\leqslant\varliminf_{n\to\infty}\|z_n\|\leqslant\varlimsup_{n\to\infty}\|z_n\|=\varlimsup_{n\to\infty}\left[\Omega(z_n)-p\Omega_1(z_n)\right]=$$

$$=\lim_{n\to\infty}\Omega(z_n)-p\lim_{n\to\infty}\Omega_1(z_n)=\bar{\Omega}-p\Omega_1(\bar{z})=\Omega(\bar{z})-p\Omega_1(\bar{z})=\|\bar{z}\|,$$

где принято во внимание также, что $\Omega(z)$, $\Omega_1(z)$ сохраняют свое значение на $\bar{Z}$. Таким образом, $\|z_n\|\to\|\bar{z}\|$ при $n\to\infty$.

Предположим теперь, что последовательность $\{z_n\}$ не сходится равномерно к $\bar{z}(s)$ на некотором отрезке $[s',s'']\subset[a,b]$, не содержащем точек разрыва функции $\bar{z}(s)$. Это означает, что найдется подпоследовательность $\{z_{n_k}\}\subset\{z_n\}$, для которой

$$\Delta_k\equiv\sup\{|z_{n_k}(s)-\bar{z}(s)|:\ s\in[s',s'']\}\geqslant\varepsilon_0=\text{const}>0\qquad(2)$$

при любом номере k. Поскольку нормы $\|z_{n_k}\|$ равномерно ограничены, то по теореме Хелли о компактности найдется подпоследовательность $\{z_r\}\subset\{z_{n_k}\}$ $r\equiv n_{k_l}$, $l=1,2,\ldots$), которая сходится к некоторой функции $z^*(s)\in V[a,b]$ в каждой точке $s\in[a,b]$ и, следовательно, П-сходится к $z^*(s)$. Вследствие П-замкнутости множества $\bar{Z}$ (см. следствие 1.4.1) функция $z^*(s)$ будет Ω-оптимальным псевдорешением. Учитывая все это, а также установленную выше сходимость $\|z_r\|\to\|\bar{z}\|=\|z^*\|$ $(r\to\infty)$, можно утверждать, что для $\{z_r\}$ выполнены условия теоремы 1.1, по которой эта последовательность будет сходиться к $z^*(s)$ равномерно на всяком отрезке, где нет точек разрыва этой функции. Но $z^*(s)$ как элемент из $\bar{Z}$ совпадает с $\bar{z}(s)$ всюду, кроме их общих точек разрыва. Поэтому из-за равномерной сходимости $\{z_r\}$ на отрезке $[s', s'']$ получим

$$\Delta_r=\sup\{|z_r(s)-\bar{z}(s)|:\ s\in[s', s'']\}=$$
$$=\sup\{|z_r(s)-z^*(s)|:\ s\in[s', s'']\}\to 0, \quad r\to\infty.$$

Это, однако, противоречит предположению (2), что и доказывает теорему.

Пусть об Ω-оптимальном псевдорешении $\bar{z}(s)$ уравнения (1) имеется априорная информация одного из следующих типов:

А) решение кусочно монотонно и, возможно, имеет конечное число обычных точек разрыва первого рода, положение которых неизвестно;

Б) решение имеет конечное число обычных точек разрыва первого рода, положение которых неизвестно, и обладает всюду на $[a, b]$, кроме точек разрыва, конечной производной класса $L_q[a, b]$ $(q\geqslant 1)$;

В) решение дифференцируемо всюду на $[a, b]$, а производная суммируема с некоторой степенью $q\geqslant 1$;

Г) решение дифференцируемо и имеет конечное число локальных экстремумов, положение которых неизвестно.

Тогда, как хорошо известно [93, 168], справедливо включение $\bar{z}(s)\in V[a, b]$ и, согласно сказанному в теореме 1, алгоритмы о.п.н., о.п.к., о.п.с.ф. гарантируют кусочно равномерную сходимость приближений в случаях А), Б), а в случаях В), Г) равномерную сходимость на всем отрезке $[a, b]$.

Рассмотрим теперь конечномерную аппроксимацию изучаемой задачи. Для этого введем семейство сеток $\Pi_N=\{a=s_1<\ldots<s_N=b\}$ при $N\geqslant 2$ на отрезке $[a, b]$. Введем также конечномерное пространство $Z_N=\mathbf{R}^N$, элементы которого $\hat{z}_N=(z_1, \ldots, z_N)$ представляют собой значения сеточной функции в узлах сетки Π_N. Следуя схеме из § 3.1, 3.4, зададим оператор P_N, действующий из $V[a, b]$ в Z_N,— оператор простого сноса функции на сетку: $P_N z=(z(s_1), \ldots, z(s_N))$ для любого $z(s)\in V[a, b]$.

Определим также для каждой сетки Π_N оператор $\bar{P}_N$: $Z_N\to V[a, b]$ как оператор восполнения значений сеточной функции между узлами. Будем считать, что он обладает следующими свойствами: 1) $\bar{P}_N$: $Z_N\to V[a, b]$ — линейный оператор; 2) для каждого $s_i\in\Pi_N$ $(i=1, \ldots, N)$ и всякого элемента $\hat{z}_N=(z_1, \ldots, z_N)\in\mathbf{R}^N$ выполнено равенство $z_N(s_i)\equiv(\bar{P}_N\hat{z}_N)(s_i)=z_i$; 3) на каждом отрезке $[s_i, s_{i+1}]$ $(i=1, \ldots, N-1)$ функция $z_N(s)$ монотонна. Такой оператор $\bar{P}_N$ будем называть *оператором монотонного восполнения.*

Нетрудно видеть, что из его свойств 2), 3) получается равенство

$$\|\bar{P}_N\hat{z}_N\|\equiv\|z_N(s)\|=|(P_N\hat{z}_N)(s_1)|+\sum_{k=1}^{N-1}\bigvee_{s_k}^{s_{k+1}}(\bar{P}_N\hat{z}_N)=|z_1|+\sum_{k=1}^{N-1}|z_{k+1}-z_k|. \quad (3)$$

Выражение, стоящее в правой части этой формулы, эквивалентно норме в $\mathbf{R}^N$, и это доказывает ограниченность при каждом фиксированном N линейного оператора $\bar{P}_N$. Таким образом, оператор $\bar{P}_N$ непрерывен из $\mathbf{R}^N$ в $V[a, b]$.

Для операторов P_N, $\bar{P}_N$ по их определению и свойствам очевидно выполнение требований 1—3 из § 3.1.

Примерами операторов $\bar{P}_N$ могут служить операторы линейной интерполяции, операторы $\bar{P}_N^*$, ставящие в соответствие вектору $\hat{z}_N \in \mathbf{R}^N$ ступенчатую функцию

$$z_N^*(s) = (\bar{P}_N^* \hat{z}_N)(s) = \{z_i\colon\ s_i \leqslant s < s_{i+1},\ i=1, \ldots, N-1;\ z_N\colon\ s = s_N\} \tag{4}$$

и ряд других. Отметим, что функция $z_N(s) = (\bar{P}_N \hat{z}_N)(s)$ может иметь точки разрыва первого рода.

Для того чтобы удовлетворить требование 4 из § 3.1 на операторы P_N, $\bar{P}_N$, будем предполагать, что функционал Ω_1 обладает в дополнение к указанным ранее свойствам еще одной особенностью:

$$\overline{\lim_{n\to\infty}}\, \Omega_1(\bar{P}_N P_N \bar{z}) \leqslant \Omega_1(\bar{z}) \quad \forall \bar{z} \in \bar{Z}.$$

Тогда, учитывая (1.1), (3), получаем

$$\begin{aligned}
&\overline{\lim_{N\to\infty}}\, \Omega_1(\bar{P}_N P_N \bar{z}) = \\
&\qquad = \overline{\lim_{N\to\infty}}\, \left\{|\bar{z}(s_1)| + \sum_{k=1}^{N-1} |\bar{z}(s_{k+1}) - \bar{z}(s_k)| + p\Omega_1(\bar{P}_N P_N \bar{z})\right\} \leqslant \\
&\qquad\qquad \leqslant \overline{\lim_{N\to\infty}}\, \left\{|\bar{z}(s_1)| + \bigvee_a^b (\bar{z}) + p\Omega_1(\bar{P}_N P_N \bar{z})\right\} \leqslant \|\bar{z}\| + p\Omega_1(\bar{z}) = \Omega(\bar{z}),
\end{aligned}$$

т. е. требование 4 из § 3.1 справедливо.

Зададим, наконец, по схеме из § 3.4 непрерывный оператор $\bar{Q}_M$, действующий из конечномерного пространства U_M размерности M в U.

Пусть далее вместо оператора A_h задан конечномерный непрерывный оператор $\hat{A}_H\colon \mathbf{R}^N \to U_M$ $(H \equiv (h, 1/M, 1/N))$ такой, что выполнено условие аппроксимации

$$\|A_h z - \bar{Q}_M \hat{A}_H P_N z\|_U \leqslant \hat{\psi}(H, \|z\|) \quad \forall z \in V[a, b]. \tag{5}$$

Мера конечномерной аппроксимации $\hat{\psi}$, удовлетворяющая обычным предположениям 3 из § 1.2, считается известной. Как будет показано в § 3, она может быть найдена для многих задач. Предположим также, что вместо величины u_σ задано ее конечномерное приближение $\hat{u}_{\sigma M} \in U_M$, подчиняющееся условию аппроксимации вида

$$\|u_\sigma - \bar{Q}_M \hat{u}_{\sigma M}\|_U \leqslant \xi(1/M, \sigma)$$

с известной функцией ξ, обладающей свойствами, описанными в § 3.4.

Итак, выполнены все условия постановки задачи о конечномерной аппроксимации уравнения (1) по схеме из § 3.1, 3.4. Поэтому для приближенного нахождения Ω-оптимальных псевдорешений этого уравнения по данным $\hat{A}_H$, $\hat{u}_{\sigma M}$, $H = (h, 1/M, 1/N)$,

$$\sigma,\ \psi(h, \|z\|),\ \hat{\psi}(H, \|z\|),\ \xi(1/M, \sigma),\ M,\ N,\ U_M,\ P_N,\ \bar{P}_N,$$

$\bar{Q}_M$ применимы конечномерные алгоритмы о.п.н., о.п.к., о.п.с.ф. Они базируются на минимизации сглаживающего функционала (3.1.5), который в нашем случае принимает вид

$$\hat{M}^{\alpha}[\hat{z}_N]=\alpha g\{|z_1|+\sum_{k=1}^{N-1}|z_{k+1}-z_k|+p\Omega_1(\bar{P}_N\,\hat{z}_N)\}+$$
$$+f(\|\bar{Q}_M\,\hat{A}_H\hat{z}_N-\bar{Q}_M\,\hat{u}_{\sigma M}\|_U),\quad \alpha>0,\quad \hat{z}_N\in\mathbf{R}^N. \quad (6)$$

Если $z_\eta(s)\equiv\bar{P}_N\hat{z}_N^{\alpha_\eta}$ $(\eta\equiv(H,\sigma))$ - - приближения, которые получаются с помощью указанных алгоритмов, то, как установлено в § 3.2—3.4, для любой последовательности $\{\eta_n\}$ такой, что $\eta_n\to0$ при $n\to\infty$, соответствующая последовательность приближений $\{z_n(s)\}$, где $z_n\equiv z_{\eta_n}$, обладает свойствами $z_n(s)\overset{\Pi}{\to}\bar{Z}$, $\Omega(z_n)\to\bar{\Omega}$ при $n\to\infty$. Рассуждая далее, как в теореме 1, можно убедиться, что эта последовательность $\{z_n(s)\}$ обладает свойством кусочно равномерной регуляризации: $z_n(s)$ равномерно сходятся к $\bar{z}(s)$ на любом отрезке $[s', s'']\subset[a, b]$, не содержащем точек разрыва функции $\bar{z}(s)$.

Замечание. Впервые задача о кусочно равномерной регуляризации рассматривалась в другой постановке в работах [9, 73], затем в [121, 122]. Функции с ограниченной вариацией использовались при решении некорректных задач в ином аспекте в работах [1, 75, 81]. При этом вопрос о кусочно равномерной регуляризации не ставился. В [67, 68, 206] изучался вариант алгоритма обобщенного метода невязки с использованием функций ограниченной вариации. Как показано в этих работах, такой алгоритм обеспечивает равномерную регуляризацию непрерывных решений.

§ 3. Некоторые примеры

Рассмотрим часто встречающиеся на практике операторные уравнения, для кусочно равномерной регуляризации которых применим подход из § 2.

Пример 1. Пусть A представляет собой интегральный оператор

$$A[x, z(s)]=\int_a^b K(x, s)\,dz(s),\quad x\in[c, d], \quad (1)$$

где $K(x, s)\in C^1(\bar{\Pi})$, $\bar{\Pi}\equiv[c, d]\times[a, b]$, а $z(s)\in V[a, b]$. Как известно, такой оператор определен на $V[a, b]$. Он будет также П-непрерывен на $V[a, b]$ в $L_2[c, d]$.

Для обоснования этого возьмем произвольную последовательность $\{z_n(s)\}\subset V[a, b]$, П-сходящуюся к $z_0(s)\in V[a, b]$. Обозначим

$$u_n(x)\equiv A[x, z_n(s)]-A[x, z_0(s)]=\int_a^b K(x, s)\,d[z_n(s)-z_0(s)].$$

Последовательность $\{u_n(x)\}$ обладает следующими свойствами:

а) она равномерно ограничена на $[c, d]$ по теореме Рисса (см. теорему 0.4.16):

$$|u_n(x)|=\left|\int_a^b K(x, s)\,d[z_n(s)-z_0(s)]\right|\leqslant\|K\|_C\cdot\|z_n-z_0\|\leqslant$$
$$\leqslant\|K\|_C(\|z_n\|+\|z_0\|)\leqslant\text{const},\quad u_n(x)\in C[c, d];$$

б) она поточечно сходится к нулю на отрезке $[c, d]$ при $n\to\infty$, что вытекает из теоремы Хелли о предельном переходе (теорема 0.4.17):

$$|u_n(x)|=\left|\int_a^b K(x, s)\,dz_n(s)-\int_a^b K(x, s)\,dz_0(s)\right|\to 0,$$

$$n\to\infty \quad \forall x\in[c, d].$$

Из этих свойств а), б) по теореме Лебега (см. следствие из теоремы 0.4.8) получим

$$\|u_n\|^2_{L_2}=\int_c^d |u_n(x)|^2\,dx\to\int_c^d 0dx=0, \quad n\to\infty,$$

и это обосновывает П-непрерывность оператора A из $V[a, b]$ в $L_2[c, d]$.

Обратимся теперь к уравнению вида (2.1):

$$A[x, z(s)]=u(x), \quad u(x)\in L_2[c, d]\equiv U. \tag{2}$$

Легко показать, что если это уравнение имеет псевдорешения на пространстве $V[a, b]$, то оно имеет их бесконечно много. Действительно, значения интеграла (1) не зависят от значений функции $z(s)$ в ее точках разрыва (см. § 0.4.п.8) и, в частности, в устранимых точках разрыва. Поэтому, «испортив» какое-либо псевдорешение уравнения (2) устранимыми точками разрыва, можно получить сколь угодно много псевдорешений, не совпадающих между собой как элементы пространства $V[a, b]$.

Нормальное псевдорешение уравнения (2), однако, не может иметь устранимых точек разрыва по лемме 1.4, а его точки разрыва первого рода обычные. Сказанное относится и к Ω-оптимальному псевдорешению, если функционал Ω_1 имеет вид (4) (см. ниже) и некоторые другие виды.

В ряде случаев Ω-оптимальные решения уравнения (2) обладают обобщенным свойством единственности. В частности, для правой части $\bar{u}(x)\in C^1[c, d]$ мы предположим выполнение этого свойства.

Будем рассматривать случай, когда оператор (1) и $\bar{u}(x)$ заданы точно $(h=0, \quad \sigma=0)$ и используется их конечномерная аппроксимация. Проверим выполнение соответствующих условий из § 2. Пусть $\Pi_N=\{s_k\colon s_k=a+\tau_N(k-1), \quad k=1, ..., N\}$ равномерная сетка на $[a, b]$ с шагом $\tau_N=(b-a)/(N-1)$ и аналогичная сетка $\Pi_M=\{x_i\colon x_i=c+\tau_M(i-1),$ $i=1, ..., M\}$, где $\tau_M=(d-c)/(M-1)$, задана на $[c, d]$. Будем полагать, что $Z_N=\mathbf{R}^N$, $U_M=\mathbf{R}^M$ — конечномерные пространства сеточных функций, определенных на сетках Π_N, Π_M соответственно. В качестве операторов $\bar{P}_N$, $\bar{Q}_M$ выберем операторы ступенчатого восполнения вида (2.4) на соответствующей сетке и определим конечномерный оператор $\hat{A}_H$ $(H=(1/M, 1/N))$, аппроксимирующий оператор (1):

$$A_H\hat{z}_N\equiv\sum_{j=1}^{N-1} K(x_i, s_j)\,\Delta z_j, \quad x_i\in\Pi_M,$$

$$s_j\in\Pi_N, \quad \Delta z_j=z_{j+1}-z_j, \quad i=1, ..., M, \quad j=1, ..., N-1.$$

Получим условие конечномерной аппроксимации.

По определению операторов $\bar{Q}_M$, P_N и по оценке $|K(x,s)-K(x_i, s_j)|\leqslant\|K'_x\|_C\cdot\tau_M+\|K'_s\|_C\cdot\tau_N$, справедливой для $x\in[x_i, x_{i+1}]$, $s\in[s_j, s_{j+1}]$ $(i=1, \ldots, M-1,\ j=1, \ldots, N-1)$, верна цепочка неравенств

$$|A[x, z(s)]-\bar{Q}_M\hat{A}_H P_N z|=$$

$$=\left|\sum_{j=1}^{N-1}\left[\int_{s_j}^{s_{j+1}} K(x, s)\,dz-K(x_i, s_j)\int_{s_j}^{s_{j+1}} dz\right]\right|=$$

$$=\left|\sum_{j=1}^{N-1}\int_{s_j}^{s_{j+1}}[K(x, s)-K(x_i, s_j)]\,dz\right|\leqslant$$

$$\leqslant\sum_{j=1}^{N-1}\|K(x, s)-K(x_i, s_j)\|_{C(\Pi_{ij})}\cdot\bigvee_{s_j}^{s_{j+1}}(z)\leqslant$$

$$\leqslant(\|K'_x\|_C\tau_M+\|K'_s\|_C\tau_N)\sum_{j=1}^{N-1}\bigvee_{s_j}^{s_{j+1}}(z)\leqslant(\|K'_x\|_C\tau_M+\|K'_s\|_C\tau_N)\,\|z\|. \qquad (3)$$

Здесь $x\in[x_i, x_{i+1}]$, $\Pi_{ij}\equiv[x_i, x_{i+1}]\times[s_j, s_{j+1}]$ и использована оценка из теоремы Рисса (см. § 0.4, п. 8):

$$\left|\int_a^b f(x)\,dz(x)\right|\leqslant\|f\|_{C[a,b]}\cdot\|z\|_{V[a,b]},\quad f\in C[a, b],\quad z\in V[a, b].$$

Из неравенств (3) легко получается соотношение

$$\|A[x, z(s)]-\bar{Q}_M\hat{A}_H P_N z\|_{L_2[c,d]}\leqslant\hat{\psi}(H, \|z\|)=$$

$$=\sqrt{d-c}\,(\|K'_x\|_C\cdot\tau_M+\|K'_s\|_C\cdot\tau_N)\,\|z\|.$$

Это вариант условия аппроксимации (2.5). Вопрос об аппроксимации правой части изучен в § 3.5.

Остановимся теперь на выборе функционала $\Omega_1(z)$ в соответствии с требованиями § 2. Можно, например, взять

$$\Omega_1(z)=z^2(b)+\int_a^b z^2(s)\,ds. \qquad (4)$$

Этот функционал определен на $V[a, b]$ как на подмножестве пространства $L_2[a, b]$, а по теореме Лебега и П-непрерывен. Выполнено также условие

$$\lim_{N\to\infty}\Omega_1(\bar{P}_N P_N\bar{z})=\Omega_1(\bar{z})\quad\forall\bar{z}\in\bar{Z},$$

вариант которого использовался в § 2.

Действительно, для оператора ступенчатого восполнения $\bar{P}_N$ имеем

$$\Omega_1(\bar{P}_N P_N z)=z^2(b)+\sum_{k=1}^{N-1}z^2(s_k)\tau_N\quad\forall z\in V[a, b],$$

где интегральная сумма справа для функции $z^2(s)\in V[a, b]$ будет сходиться при $N\to\infty$ в силу ее суммируемости к интегралу из формулы (4).

Таким образом, из сказанного ясно, что для кусочно равномерной регуляризации уравнений вида (2) можно использовать алгоритмы к.о.п.н., к.о.п.к., к.о.п.с.ф., как обосновано в § 2.

Пример 2. Пусть

$$A[x, z(s)] = \int_a^b K(x, s, z(s))\, ds, \quad x \in [c, d], \ z(s) \in V[a, b], \tag{5}$$

где функция $K(x, s, z) \in C(\Pi_0)$ $(\Pi_0 = [c, d] \times [a, b] \times \mathbf{R})$ удовлетворяет условиям из примера 3.5.1. Из неравенства

$$|K(x, s, z_1(s)) - K(x, s, z_2(s))| \leqslant \xi_3(z_0)\, |z_1(s) - z_2(s)|,$$

$$z_0 \equiv \max\{\|z_1\|, \ \|z_2\|\},$$

полученного, как в примере 3.5.1, по теореме Лебега можно сделать вывод о том, что оператор (5) П-непрерывен из $V[a, b]$ в $L_2[c, d]$. Конечномерная аппроксимация оператора (5) проводится по той же схеме, что и в примере 1, если положить

$$\hat{A}_H \hat{z} = \sum_{j=1}^{N} K(x_i, s_j, z_j)\, \tau_N, \quad x_i \in \Pi_M, \quad s_j \in \Pi_N,$$

$$i = 1, \ldots, M, \quad j = 1, \ldots, N, \quad \hat{z} = (z_1, \ldots, z_N) \in Z_N.$$

Тогда по аналогии с примером 3.5.1, используя технику примера 1, получаем условия аппроксимации

$$\|A[x, z(s)] - \bar{Q}_M \hat{A}_H P_N z\|_{L_2[c, d]} \leqslant \hat{\psi}(H, \|z\|) =$$

$$= \sqrt{d-c}\,\{[\xi_0(\|z\|) + \xi_2(\|z\|)(b-a)/2 + \xi_3(\|z\|) \cdot \|z\|]\, \tau_N + \xi_1(\|z\|)(b-a)\, \tau_M\}.$$

Можно также привести и другие примеры задач, для которых применима теория § 2, а также другие способы конечномерной аппроксимации задач из примеров 1, 2.

§ 4. О реализации алгоритмов кусочно равномерной регуляризации

Алгоритмы кусочно равномерной регуляризации, предложенные в § 2, основаны в своем конечномерном варианте на решении задачи минимизации сглаживающего функционала: найти элемент $\hat{z}^\alpha \in \mathbf{R}^N$ такой, для которого

$$\hat{M}^\alpha[\hat{z}^\alpha] = \inf\{\hat{M}^\alpha[\hat{z}]: \ \hat{z} \in \mathbf{R}^N\}, \quad \hat{z} = (z_1, \ldots, z_N). \tag{1}$$

При этом

$$\begin{gathered} \hat{M}^\alpha[\hat{z}] = \alpha\{|z_1| + \sum_{k=1}^{N-1} |z_{k+1} - z_k|\} + \Phi_\alpha(\hat{z}), \\ \Phi_\alpha(\hat{z}) \equiv \alpha p \Omega_1(\bar{P}_N \hat{z}) + f(\|\bar{Q}_M \hat{A}_H \hat{z} - \bar{Q}_M \hat{u}_{\sigma M}\|_U) \end{gathered} \tag{2}$$

(см. (2.6)). Минимизация негладкого конечномерного функционала (2) связана с определенными трудностями. В данном параграфе мы рассмотрим методику, которая позволяет вместо задачи безусловной минимизации (1) с негладким функционалом (2) решать другую задачу, но уже с некоторым специальным дифференцируемым функционалом и с простыми ограничениями.

Определим множество $\hat{Z}\uparrow$ векторов из пространства $\mathbf{R}^N$, координаты которых подчинены неравенствам $z_k \leqslant z_{k+1}$ $(k=1, \ldots, N-1)$. Введем векторы $\hat{v}$, $\hat{w} \in \mathbf{R}^N$ по правилам:

$$\hat{v}=(v_1, \ldots, v_N), \quad \hat{w}=(w_1, \ldots, w_N); \quad v_1=\{z_1: z_1 \geqslant 0; \; 0: z_1<0\},$$

$$w_1=\{0: z_1 \geqslant 0; \; -z_1: z_1<0\}; \quad v_i=\sum_{k=1}^{i-1} |z_{k+1}-z_k|,$$

$$w_i=v_i-z_i, \quad i=2, \ldots, N.$$

Ясно, что $\hat{z}=\hat{v}-\hat{w}$ и при этом $\hat{v}$, $\hat{w} \in \hat{Z}\uparrow$. Такое представление вектора $\hat{z}$ дает возможность сделать замену переменных:

$$\xi_1=v_1 \geqslant 0, \quad \xi_{k+1}=v_{k+1}-v_k \geqslant 0, \quad \eta_1=w_1 \geqslant 0,$$

$$\eta_{k+1}=w_{k+1}-w_k, \quad k=1, \ldots, N-1,$$

при которой выполнено равенство

$$z_k=\sum_{l=1}^{k} (\xi_l-\eta_l), \quad k=1, \ldots, N. \tag{3}$$

После такой замены переменных функционал (2) принимает вид

$$\hat{M}^\alpha[\hat{\xi}, \hat{\eta}]=\alpha \sum_{l=1}^{N} |\xi_l-\eta_l|+\Phi(\hat{\xi}-\hat{\eta}),$$

где $\hat{\xi}=(\xi_1, \ldots, \xi_N) \in \mathbf{R}^N_+$, $\hat{\eta}=(\eta_1, \ldots, \eta_N) \in \mathbf{R}^N_+$, а новый функционал Φ определен с учетом (3) равенством $\Phi(\hat{\xi}-\hat{\eta})=\Phi_0(\hat{z})$. При этом задача (1) сводится к следующей конечномерной задаче минимизации с ограничениями: найти элементы $\hat{\xi}^*$, $\hat{\eta}^* \in \mathbf{R}^N_+$, для которых

$$\hat{M}^\alpha[\hat{\xi}^*, \hat{\eta}^*]=\inf\{\hat{M}^\alpha[\hat{\xi}, \hat{\eta}]: \hat{\xi}, \; \hat{\eta} \in \mathbf{R}^N_+\}. \tag{4}$$

Рассмотрим более подробно эту задачу. Оказывается, что ее можно свести к другой задаче: найти такие векторы $\hat{a}^*$, $\hat{b}^* \in \mathbf{R}^N_+$, что

$$\hat{L}^\alpha[\hat{a}^*, \hat{b}^*]=\inf\{\hat{L}^\alpha[\hat{a}, \hat{b}]: \hat{a}, \; \hat{b} \in \mathbf{R}^N_+\}. \tag{5}$$

В этой новой задаче функционал $\hat{L}^\alpha[\hat{a}, \hat{b}]$ имеет вид

$$\hat{L}^\alpha[\hat{a}, \hat{b}]=\alpha \sum_{k=1}^{N} (a_k+b_k)+\Phi(\hat{a}-\hat{b}) \equiv \alpha\hat{\Omega}(\hat{a}, \hat{b})+\Phi(\hat{a}-\hat{b}).$$

Такой переход обычно осуществляется в *методе наименьших модулей* (см., например, [83, 165]). Связь задач (4) и (5) устанавливает

Теорема 1. *Если* $(\hat{a}^*, \hat{b}^*)$ — *решение задачи* (5), *то вектор* $(\hat{\xi}^*, \hat{\eta}^*)=(\hat{a}^*, \hat{b}^*)$ *будет также решением задачи* (4).

Доказательство. Прежде всего заметим, что решение задачи (5) обладает характерной особенностью: для каждого $k=1, \ldots, N$ числа a^*_k, b^*_k не могут быть одновременно положительными. Действительно, если это не так и для $k=k_0$ оказывается, что $a^*_{k_0}>0$, $b^*_{k_0}>0$, то можно ввести положительное число $\lambda\equiv\min(a^*_{k_0}, b^*_{k_0})$. С его помощью можно также определить элементы $\hat{a}^{**}$, $\hat{b}^{**}\in\mathbf{R}^N_+$ по правилу: $a^{**}_k=a^*_k$, $b^{**}_k=b^*_k$ при $k\neq k_0$, $a^{**}_{k_0}=a^*_{k_0}-\lambda$, $b^{**}_{k_0}=b^*_{k_0}-\lambda$. Тогда очевидно соотношение

$$\hat{L}^\alpha[\hat{a}^{**}, \hat{b}^{**}]=\hat{L}^\alpha[\hat{a}^*, \hat{b}^*]-2\alpha\lambda<\hat{L}^\alpha[\hat{a}^*, \hat{b}^*],$$

которое противоречит экстремальности элемента $(\hat{a}^*, \hat{b}^*)$ как решения задачи (5).

Установленное таким образом характерное свойство решений задачи (5) приводит к равенству

$$\sum_{k=1}^N (a^*_k+b^*_k)=\sum_{k=1}^N |a^*_k-b^*_k|, \tag{6}$$

из которого следует, что $\hat{L}^\alpha[\hat{a}^*, \hat{b}^*]=\hat{M}^\alpha[\hat{a}^*, \hat{b}^*]$. Если теперь предположить, что $(\hat{a}^*, \hat{b}^*)$ не является решением задачи (4), то найдутся такие векторы $\xi, \eta\in\mathbf{R}^N_+$, для которых выполнено неравенство $\hat{M}^\alpha[\hat{a}^*, \hat{b}^*]>\hat{M}^\alpha[\xi, \eta]$. Введем новые векторы $\hat{\xi}^{**}$, $\hat{\eta}^{**}\in\mathbf{R}^N_+$ по правилам:

$$\xi^{**}_k=\{\xi_k-\eta_k\colon\ \xi_k>\eta_k;\ 0\colon\ \xi_k\leqslant\eta_k\},$$

$$\eta^{**}_k=\{0\colon\ \xi_k>\eta_k;\ \eta_k-\xi_k\colon\ \xi_k\leqslant\eta_k\}.$$

Тогда, учитывая вид функционалов $\hat{M}^\alpha$, $\hat{L}^\alpha$, получаем

$$\hat{M}^\alpha[\hat{a}^*, \hat{b}^*]=\hat{L}^\alpha[\hat{a}^*, \hat{b}^*]>\hat{M}^\alpha[\hat{\xi}, \hat{\eta}]=\hat{L}^\alpha[\hat{\xi}^{**}, \hat{\eta}^{**}].$$

Это неравенство противоречит экстремальности величины $(\hat{a}^*, \hat{b}^*)$ как решения задачи (5). Таким образом, для любых $\hat{\xi}$, $\hat{\eta}\in\mathbf{R}^N_+$ должно выполняться неравенство $\hat{M}^\alpha[\hat{a}^*, \hat{b}^*]\leqslant\hat{M}^\alpha[\hat{\xi}, \hat{\eta}]$, что и доказывает теорему.

Из теоремы 1 ясно, что для нахождения какой-либо экстремали задачи (1) достаточно решить задачу (5) с последующей заменой переменных по формуле (3).

Теорема 2. *При выполнении предположений из § 2 на конечномерные величины* $\bar{P}_N$, $\bar{Q}_M$, Ω_1, $\hat{A}_H$, *а также функцию* $f(x)$, *задача* (5) *разрешима.*

Действительно, по условиям теоремы функционал $\Phi(\hat{a}-\hat{b})$ непрерывен на множестве $\mathbf{R}^N_+\times\mathbf{R}^N_+$ так же, как и функционал $\hat{\Omega}(\hat{a}, \hat{b})$. Кроме того, функционал $\hat{\Omega}(\hat{a}, \hat{b})$ обладает вытекающим из оценки $0\leqslant a_k+b_k\leqslant\hat{\Omega}(\hat{a}, \hat{b})$ $(a_k\geqslant 0,\ b_k\geqslant 0)$ свойством компактности в $\mathbf{R}^N_+\times\mathbf{R}^N_+$ множеств $\hat{\Omega}_C=\{\hat{a}, \hat{b}\in\mathbf{R}^N_+\colon\ \hat{\Omega}(\hat{a}, \hat{b})\leqslant C\}$ $(C\geqslant 0)$. Тогда задача (5) разрешима по общей теореме 1.4.2.

Во многих случаях задача (5) представляется гораздо более простой, чем задача (1). Это относится особенно к случаю решения линейных операторных уравнений (2.1), когда задача (5) оказывается задачей квадратичного программирования на множестве векторов с неотрицательными координатами. Известны эффективные конечношаговые численные методы решения подобных задач [173].

При использовании для кусочно равномерной регуляризации уравнения (2.1) алгоритмов о.п.н., о.п.к., о.п.с.ф. вопросы численной реализации решаются так же, как и в общем случае (см. § 1.14). Поэтому достаточно

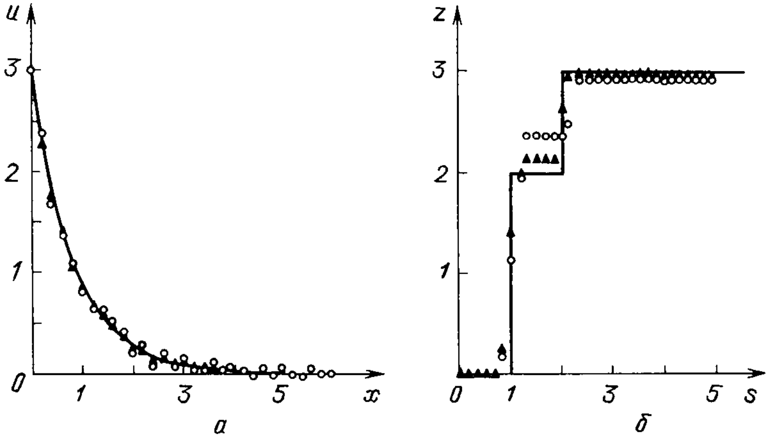

Рис. 17. Кусочно равномерная регуляризация уравнения (3.2) с ядром $K(x, s)=e^{-xs}$ и правой частью $\bar{u}=2e^{-x}+e^{-2x}$ по алгоритму о.п.н.: $а$ -- точная и приближенные правые части; $б$ -- точное и приближенное решения. Точные величины —; приближенные величины: $\sigma=10^{-2}$ ○○○; $\sigma=10^{-3}$ ▲▲▲

при произвольном $\alpha>0$ уметь находить одно решение $\hat{z}^{\alpha}$ задачи (1). Это можно делать путем решения задачи (5), как указано выше.

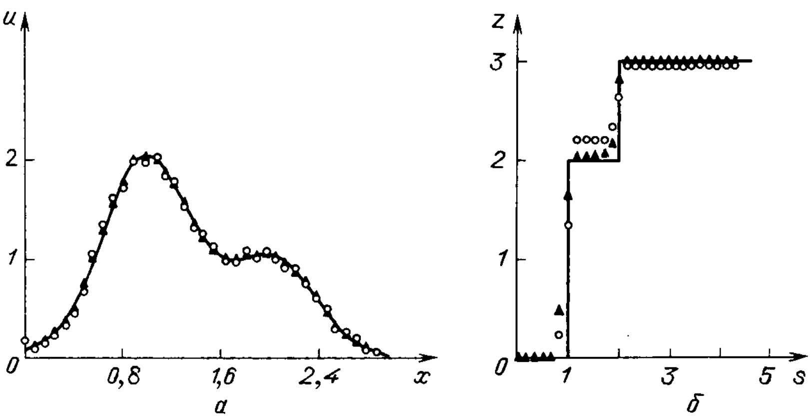

Рис. 18. Кусочно равномерная регуляризация уравнения (3.2) с ядром $K(x, s)=\exp\{-4(x-s)^2\}$ и правой частью $\bar{u}=2\exp\{-4(x-1)^2\}+\exp\{-4(x-2)^2\}$

В заключение главы приведем примеры модельных расчетов по кусочно равномерной регуляризации нормальных решений уравнения вида (3.2) по схеме § 2, 3. В этих примерах минимизация сглаживающего функционала (1) с $p=0$ проводилась по методике данного параграфа: решалась задача (5) и по полученному ее решению $\hat{\xi}^\alpha=\hat{a}^\alpha$, $\hat{\eta}^\alpha=\hat{b}^\alpha$ по формуле (3) вычислялся элемент $\hat{z}^\alpha$, минимизирующий сглаживающий функционал в $\mathbf{R}^N$.

Результаты некоторых расчетов приведены на рис. 17—19. При этом на них показаны точные величины $\bar{u}$, $\bar{z}$, а также вектор приближенной правой части $\hat{u}_{\sigma M}$ уравнения (3.2). Там же представлен вектор $\hat{z}^{\alpha_\eta}$, полученный по алгоритму кусочно равномерной регуля-

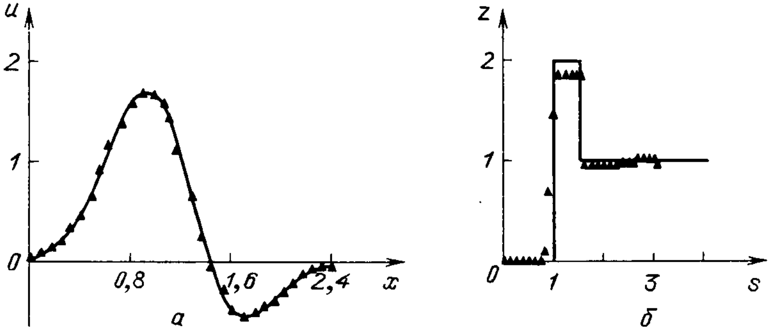

Рис. 19. Кусочно равномерная регуляризация уравнения (3.2) с ядром $K(x, s)=$ $=\exp\{-4(x-s)^2\}$ и правой частью $\bar{u}=2\exp\{-4(x-1)^2\}-\exp\{-4(x-3/2)^2\}$

ризации (к.м.о.п.н.) для различных уровней возмущения σ правой части уравнения.

Рис. 17 относится к решению уравнения (3.2) с ядром $K(x, s)=$ $=\exp\{-xs\}$, а рис. 18, 19—к уравнению с ядром $K(x, s)=$ $=\exp\{-4(x-s)^2\}$. Заметим, что хорошо известна высокая чувствительность приближенных решений уравнений вида (3.2) с ядром первого типа к погрешностям правой части [114]. Это и видно из рис. 17.

Замечание. Численные методы решения некорректных задач на множествах монотонных, выпуклых и некоторых других функций описаны в [57, 68, 206]. В [206] приведены тексты программ на языке Фортран, реализующие эти методы.

ГЛАВА 5

ИСПОЛЬЗОВАНИЕ ОБОБЩЕННОГО ПРИНЦИПА НЕВЯЗКИ ДЛЯ РЕШЕНИЯ НЕКОТОРЫХ ЗАДАЧ ЛИНЕЙНОЙ АЛГЕБРЫ

§ 1. Основная задача

В данной главе будут изучаться некоторые устойчивые методы решения систем линейных алгебраических уравнений (СЛАУ) вида

$$\bar{A}z=\bar{u}, \quad z\in\mathbf{R}^n, \quad \bar{u}\in\mathbf{R}^m. \tag{1}$$

При этом предполагается, что $\bar{A}$ — ненулевая действительная матрица размера $m\times n$ и, кроме того, $\bar{u}\neq 0$. Эта система может не быть разрешимой в классическом смысле. Однако она всегда имеет непустое множество Z^* псевдорешений, т. е. таких векторов $z^*\in\mathbf{R}^n$, для которых

$$\|\bar{A}z^*-\bar{u}\|=\inf\{\|\bar{A}z-\bar{u}\|:\ z\in\mathbf{R}^n\}$$

(все нормы, используемые в данной главе, считаются евклидовыми). Для совместной системы (1) эти псевдорешения представляют собой обычные решения.

Известно, что для СЛАУ (1) существует единственное нормальное псевдорешение $\bar{z}$, т. е. такое псевдорешение, для которого выполнено равенство

$$\|\bar{z}\|=\inf\{\|z\|:\ z\in Z^*\}.$$

Для разрешимой системы (1) такое нормальное псевдорешение называют *нормальным решением* [197, 202]. В дальнейшем будем считать выполненным неравенство $\bar{z}\neq 0$.

Предположим далее, что вместо точных данных $(\bar{A}, \bar{u})$ системы (1) заданы их приближения (A_h, u_σ). Здесь A_h — приближенная матрица размера $m\times n$, удовлетворяющая условию $\|A_h-\bar{A}\|\leqslant h$. Для приближенной правой части u_σ системы выполнено неравенство $\|u_\sigma-\bar{u}\|\leqslant\sigma$.

Легко убедиться в том, что задача нахождения нормального псевдорешения в общем случае неустойчива по отношению к возмущениям матрицы. Для этого достаточно рассмотреть

Пример 1. Пусть

$$\bar{A}=\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad \bar{u}=\begin{bmatrix}1\\1\end{bmatrix},\quad A_h=\begin{bmatrix}1&0\\0&h\end{bmatrix},\quad u_\sigma=\bar{u},\quad \sigma=0,\quad z=\begin{bmatrix}x\\y\end{bmatrix}.$$

Тогда, вычисляя непосредственно по определению нормальное псевдорешение такой системы (1), получаем $\bar{z}=(1,\ 0)^T$. Система $A_h z=u_\sigma$ с возмущенными данными имеет единственное решение $\bar{z}_h=(1,\ 1/h)^T$. Таким образом, при $h\to 0$ отсутствует сходимость величин $\bar{z}_h$ к $\bar{z}$, что и означает неустойчивость нормального псевдорешения данной задачи.

Основная задача данной главы заключается в построении по приближенным данным $(A_h,\ u_\sigma,\ h,\ \sigma)$ устойчивого в $\mathbf{R}^n$ приближения z_δ $(\delta\equiv(h,\ \sigma))$ к нормальному псевдорешению $\bar{z}$ системы (1): $\|z_\delta-\bar{z}\|\to 0$ при $\delta\to 0$.

В принципе данная постановка задачи представляет собой частный случай общей постановки из § 2.1, в которой следует принять $Z=D=\mathbf{R}^n$, $U=\mathbf{R}^m$, $\Omega(z)=\|z\|^2$, $\psi(\delta,\ \Omega)=\sigma+h\,\|z\|$, $\bar{Z}=\{\bar{z}\}$. Поэтому для решения основной задачи этой главы могут быть использованы алгоритмы, разработанные в гл. 2. При этом вследствие линейности и конечномерности задачи (1) алгоритмы приобретают ряд дополнительных свойств и особенностей. Рассмотрим эти особенности на примере алгоритма *обобщенного принципа невязки* (о.п.н.) для нахождения нормального псевдорешения $\bar{z}$.

Алгоритм о.п.н. для решения основной задачи удобно основывать на использовании сглаживающего функционала вида

$$M^\alpha[z]=\alpha\|z\|^2+\|A_h z-u_\sigma\|^2,\quad z\in\mathbf{R}^n,\quad \alpha>0. \tag{2}$$

Тогда задача минимизации сглаживающего функционала: найти при фиксированном δ и данном $\alpha>0$ элемент z^α, для которого

$$M^\alpha[z^\alpha]=\inf\{M^\alpha[z]\colon z\in\mathbf{R}^n\}, \tag{3}$$

будет иметь единственное решение. Это следует, например, из приведенного в § 0.5 примера 4 или из леммы 1.11.1, так как сглаживающий функционал представляет собой сумму сильно выпуклого в $\mathbf{R}^n$ функционала $\|z\|^2$ и выпуклого функционала невязки $\|A_h z-u_\sigma\|^2$. Кроме того, необходимым и достаточным условием того, что z^α является решением экстремальной задачи (3) с сильно выпуклым и непрерывно дифференцируемым в $\mathbf{R}^n$ функционалом (2), есть требование $\operatorname{grad} M^\alpha[z^\alpha]=0$. Оно эквивалентно уравнению

$$\alpha z^\alpha+A_h^T A_h z^\alpha=A_h^T u_\sigma \tag{4}$$

(см., например, [202] или пример 0.5.4). Это операторное уравнение — система линейных алгебраических уравнений для нахождения z^α — хорошо изучено и эффективно решается.

Введем теперь обобщенную невязку

$$\rho(\alpha)=\|A_h z^\alpha-u_\sigma\|^2-(\lambda_\delta+\sigma+h\,\|z^\alpha\|)^2,\quad \alpha>0, \tag{5}$$

используемую для выбора параметра регуляризации. Здесь число

$$\lambda_\delta=\inf\{\|A_h z-u_\sigma\|+\sigma+h\,\|z\|\colon z\in\mathbf{R}^n\}$$

есть обобщенная мера несовместности СЛАУ (1), введенная в [39, 105]. Вследствие единственности решения задачи (3) обобщенная невязка будет непрерывной при $\alpha>0$ (см. § 1.6). Ее предельные значения $\rho_0=\rho(+0)$, $\rho_\infty=\rho(+\infty)$ удовлетворяют при достаточно малых σ, h неравенствам $\rho_0<0$, $\rho_\infty>0$. Это следует из теоремы 1.7.3, так как мера аппроксимации $\psi(\delta, \Omega)=\sigma+h\Omega^{1/2}$ ($h\neq 0$) монотонно возрастает по второму аргументу и выполнено условие $\bar{z}\neq 0$, которое в обозначениях гл. 1 может быть записано в виде $Z_0\bigcap Z^*=\varnothing$.

Из теоремы 1.11.1 следует, что уравнение $\rho(\alpha)=0$, используемое для выбора параметра регуляризации α_δ по алгоритму о.п.н., имеет единственное решение $\alpha_\delta>0$. Сама функция $\rho(\alpha)$ монотонно возрастает при $\alpha>0$ (см. лемму 2.4.2).

Таким образом, алгоритм о.п.н. порождает при фиксированном $\delta=(h, \sigma)$ единственное приближенное решение $z_\delta\equiv z^{\alpha_\delta}$ системы (1) которое по теореме 1.7.5 сходится к $\bar{z}$ при $\delta\to 0$.

Если имеется дополнительная информация о совместности системы (1), то в формуле (5) можно положить $\lambda_\delta=0$ (см. § 2.3).

Алгоритм о.п.н. для решения основной задачи использует данные $\{A_h, u_\sigma, h, \sigma\}$. Можно, однако, предложить его модификацию, которая не использует явно числа σ-погрешности задания правой части СЛАУ (см. [129]).

Для удобства формулировок введем нормированные пространства матриц размерности $m\times n$, $n\times m$, $m\times m$, $n\times n$ с евклидовыми нормами $\|\cdot\|$, $\|\cdot\|_*$, $\|\cdot\|_m$, $\|\cdot\|_n$ и будем обозначать эти пространства $\mathfrak{A}$, $\mathfrak{A}^*$, $\mathfrak{A}_m$, $\mathfrak{A}_n$ соответственно. Известно, что нормальное псевдорешение $\bar{z}$ задачи (1) выражается через точные ее данные $(\bar{A}, \bar{u})$ следующим образом: $\bar{z}=\bar{A}^+\bar{u}$. Фигурирующая в этом равенстве псевдообратная матрица $\bar{A}^+$ представляет собой решение экстремальной задачи: найти такие матрицы $\tilde{Z}\in\mathfrak{A}^*$, для которых

$$\|\bar{A}\tilde{Z}-E\|_m=\inf\{\|\bar{A}Z-E\|_m\colon Z\in\mathfrak{A}^*\}. \tag{6}$$

Здесь $E\in\mathfrak{A}_m$ — единичная матрица. Эта задача может иметь не единственное решение. Легко убедиться, например, что кроме $\tilde{Z}=\bar{A}^+$ задача имеет также решения вида $\tilde{Z}=\bar{A}^+ +q(I-\bar{A}^+\bar{A})$, где $I\in\mathfrak{A}_n$ — единичная матрица, а q — произвольная константа. При этом единственным нормальным решением задачи (6) является $\bar{A}^+$: если $\mathfrak{A}_0^*$ — множество решений задачи (6), то

$$\|\bar{A}^+\|_*=\inf\{\|\tilde{Z}\|_*\colon \tilde{Z}\in\mathfrak{A}_0^*\}. \tag{7}$$

Таким образом, экстремальная задача (6), (7) о нахождении псевдообратной матрицы $\bar{A}^+$ является частным случаем задачи из гл. 1. Поэтому для построения регуляризованных приближений к этой псевдообратной матрице можно применить, например, обобщенный принцип невязки для экстремальных задач (см. § 1.7). Он может быть основан на использовании сглаживающего функционала вида

$$M^\alpha[Z]=\alpha\|Z\|_*^2+\|A_hZ-E\|_m^2,\quad \alpha>0,\quad Z\in\mathfrak{A}^*. \tag{8}$$

Этот функционал имеет при каждом $\alpha>0$ единственную экстремаль

$$Z^\alpha=(\alpha I+A_h^TA_h)^{-1}A_h^T,$$

реализующую его точную нижнюю грань в $\mathfrak{A}^*$. Поэтому можно ввести непрерывную монотонно возрастающую при $\alpha>0$ обобщенную невязку

$$\rho(\alpha)=\|A_h Z^\alpha - E\|_m^2-(\Lambda_h+h\|Z^\alpha\|_*)^2,$$

в которой

$$\Lambda_h \equiv \inf\{\|A_h Z-E\|_m+h\|Z\|_*: Z\in\mathfrak{A}^*\}.$$

Из теории, данной в § 1.7, следует, что уравнение $\rho(\alpha)=0$ имеет единственное решение $\alpha(h)>0$, если погрешность $h>0$ достаточно мала. Взяв в качестве приближения к $\bar{A}^+$ матрицу $Z^{\alpha(h)}$, получим по теореме 1.7.5 сходимость $Z^{\alpha(h)}\to\bar{A}^+$ при $h\to 0$.

Возможны и другие подходы к построению регуляризованных приближений к $\bar{A}^+$ (см., например, [28, 147, 164]).

Регуляризованные приближения $Z^{\alpha(h)}$ можно использовать для построения приближений к $\bar{z}$: $z_\delta=Z^{\alpha(h)}u_\sigma$. Тогда из неравенства

$$\|z_\delta-\bar{z}\|\leqslant\|Z^{\alpha(h)}-\bar{A}^+\|_*\cdot\|u_\sigma\|+\|\bar{A}^+\|_*\cdot\|u_\sigma-\bar{u}\|$$

получается сходимость $z_\delta\to\bar{z}$ при $\delta=(h,\sigma)\to 0$.

Изложенный алгоритм нахождения приближений к нормальному псевдорешению СЛАУ (1) не использует величины σ, но существенно использует величину h. Такая возможность открывается вследствие специфики задачи (1) (нормальная разрешимость оператора $\bar{A}$). Из результатов работы [18] следует, что построить регуляризующий алгоритм для решения СЛАУ общего вида (1) без использования погрешности h невозможно.

Для алгоритмов устойчивого решения СЛАУ (1), основанных на использовании сглаживающего функционала (2) (или (8)) (и, в частности, для алгоритма о.п.н.), имеется принципиальное различие случаев совместной и несовместной СЛАУ (1). Это различие, отмечавшееся, например, в [28, 55], связано с порядком наилучшей достижимой точности приближенного решения.

Чтобы разобраться в этом, рассмотрим сначала пример несовместной системы. Пусть данные $\bar{A}$, $\bar{u}$, A_h, u_σ определены, как в примере 1. Тогда функционал (2) примет вид

$$M^\alpha[z]=M^\alpha(x,y)=\alpha(x^2+y^2)+(x-1)^2+(hy-1)^2.$$

Решая теперь задачу (3), получаем

$$z^\alpha=(x^\alpha,y^\alpha)^T,\quad x^\alpha=1/(1+\alpha),\quad y^\alpha=h/(\alpha+h^2),$$

откуда

$$\|z^\alpha-\bar{z}\|^2=\alpha^2/(1+\alpha)^2+h^2/(\alpha+h^2)^2\equiv p^2(\alpha).$$

Нетрудно убедиться, исследуя функцию $p(\alpha)$, что $\min p(\alpha)\asymp\sqrt{h}$. Итак, в данном примере при любом выборе параметра регуляризации точность приближенного решения не может быть по порядку лучше чем $\sqrt{h}$.

С другой стороны, оказывается, что для совместных СЛАУ (1) справедлива

Теорема 1. *При поиске нормального решения $\bar{z}$ совместной системы* (1) *на основе метода регуляризации с использованием сглаживающего функционала* (2) *существует такой выбор параметра регуляризации* $\alpha=\alpha(h)$, *что при* $0\leqslant h<h_0=\mathrm{const}$ *и при любом* $\sigma\geqslant 0$ *выполнена оценка*

$$\|z^{\alpha(h)}-\bar{z}\|\leqslant\frac{\|\bar{A}^+\|_*\cdot[\sigma+h(1+\sqrt{2})\|\bar{z}\|]}{1-h\|\bar{A}^+\|_*}=O(h+\sigma). \tag{9}$$

Эта оценка точна по порядку входящих в нее величин σ, h, $\|\bar{A}^+\|_*$.

Данный результат установлен в работе [28]. Порядок точности оценки (9) изучался также в [55, 163, 164].

Таким образом, для разрешимых СЛАУ метод регуляризации (при надлежащем выборе параметра α) может давать оптимальный порядок точности $O(h+\sigma)$ приближенного решения. Более подробно об оптимальном порядке точности сказано в § 5.

Возвратимся к алгоритму о.п.н. Рассмотренный пример показывает, что о.п.н. в случае несовместных СЛАУ может не давать оптимальный порядок точности. Покажем, что для совместных систем (1) порядок точности приближенного решения, полученного по алгоритму о.п.н., будет оптимальным. При этом будем использовать модифицированный алгоритм о.п.н., в котором параметр регуляризации α_δ выбирается как единственный корень уравнения с монотонно возрастающей функцией

$$\rho(\alpha)=\|A_h z^\alpha-u_\sigma\|-C(\sigma+h\|z^\alpha\|)=0,\quad C>1 \tag{10}$$

(ср. (5)). В § 2.3 обоснована возможность такого выбора параметра регуляризации, а также сходимость $z^{\alpha_\delta}\to\bar{z}$ при $\delta\to 0$. При этом $\|z^{\alpha_\delta}\|\leqslant\|\bar{z}\|$.

Теорема 2. *Если* $\bar{u}\in\operatorname{Im}\bar{A}$, *т. е. СЛАУ* (1) *разрешима, то алгоритм м.о.п.н., основанный на выборе* α_δ *из уравнения* (10), *имеет оптимальный порядок точности.*

Для доказательства используем леммы.

Лемма 1 [28]. *Если* $\bar{u}\in\operatorname{Im}\bar{A}$, $b=\mathrm{const}>\|\bar{A}^+\|_*/(\|\bar{A}^+\|_*-h)$ *и число* $\alpha(\delta)$ *удовлетворяет уравнению*

$$\beta(\alpha)\equiv\|A_h z^\alpha-u_\sigma\|=b(\sigma+h\|\bar{z}\|),$$

то $\|z^{\alpha(\delta)}-\bar{z}\|=O(\sigma+h)$.

Лемма 2. *Если* $0<\alpha_1\leqslant\alpha\leqslant\alpha_2$, *то справедливо неравенство* $\|z^{\alpha_2}-z^\alpha\|\leqslant\|z^{\alpha_2}-z^{\alpha_1}\|$.

Доказательство. Используя сингулярное разложение $A_h=URV^T$ матрицы A_h (см. § 0.6), можно найти из уравнения (4) элемент $z^\alpha=V(\alpha I+R^TR)^{-1}R^TU^Tu_\sigma$. Тогда, используя обозначение $w\equiv U^Tu_\sigma$ и учитывая ортогональность матрицы V, получаем

$$\|z^{\alpha_2}-z^\alpha\|^2=\|V[\alpha_2 I+R^TR)^{-1}R^Tw-(\alpha I+R^TR)^{-1}R^Tw]\|^2=$$

$$=\|(\alpha_2 I+R^TR)^{-1}(\alpha I+R^TR)^{-1}(\alpha I+R^TR-\alpha_2 I-R^TR)R^Tw\|^2=$$

$$=(\alpha-\alpha_2)^2\cdot\|(\alpha_2 I+R^TR)^{-1}(\alpha I+R^TR)^{-1}R^Tw\|^2.$$

Запишем последнее равенство в развернутом виде, учитывая, что $R=\mathrm{diag}(\rho_1, \ldots, \rho_M)\in\mathfrak{A}$, $M\equiv\min(m, n)$, $w=(w_1, \ldots, w_m)$. Тогда справедлива оценка

$$\| z^{\alpha_2}-z^{\alpha}\|^2=(\alpha_2-\alpha)^2\sum_{k=1}^{M}\frac{\rho_k^2 w_k^2}{(\alpha_2+\rho_k^2)^2(\alpha+\rho_k^2)^2}\leqslant$$

$$\leqslant(\alpha_2-\alpha_1)^2\sum_{k=1}^{M}\frac{\rho_k^2 w_k^2}{(\alpha_2+\rho_k^2)^2(\alpha_1+\rho_k^2)^2}=\| z^{\alpha_2}-z^{\alpha_1}\|,$$

которая и доказывает лемму.

Доказательство теоремы 2. Пусть ε $(0<\varepsilon<1)$ — такое число, что $C(1-\varepsilon)>1$. Выберем столь малые числа σ_0, h_0, что при $0\leqslant\sigma<\sigma_0$, $0\leqslant h<h_0$ будут выполнены неравенства $\| z^{\alpha_\delta}\|\geqslant(1-\varepsilon)\|\bar z\|$, $C(1-\varepsilon)>\|\bar A^+\|_*/(\|\bar A^+\|_*-h)$. Такие числа σ_0, h_0 существуют, ибо в противном случае нашлась бы последовательность $\delta_n=(\sigma_n, h_n)\to 0$ при $n\to\infty$, для которой были бы выполнены противоположные неравенства:

$$\| z^{\alpha_{\delta_n}}\|<(1-\varepsilon)\|\bar z\|,\quad C(1-\varepsilon)\leqslant\|\bar A^+\|_*/(\|\bar A^+\|_*-h_n).$$

Переходя в этих неравенствах к пределу при $n\to\infty$, можно было бы получить с учетом сходимости $z^{\alpha_{\delta_n}}\to\bar z$, что $\|\bar z\|\leqslant(1-\varepsilon)\|\bar z\|$, $C(1-\varepsilon)\leqslant 1$. Это, однако, противоречит требованию на число ε.

Итак, пусть $0\leqslant\sigma<\sigma_0$, $0\leqslant h<h_0$. Тогда выполнено неравенство

$$C(1-\varepsilon)(\sigma+h\|\bar z\|)\leqslant C(\sigma+h(1-\varepsilon)\|\bar z\|)\leqslant C(\sigma+h\| z^{\alpha_\delta}\|)\leqslant C(\sigma+h\|\bar z\|),$$

в котором учтено, что $\| z^{\alpha_\delta}\|\leqslant\|\bar z\|$. Из этого неравенства в силу возрастания функции $\beta(\alpha)$ следует, что если выбрать числа $\alpha_1(\delta)$, $\alpha_2(\delta)$ соответственно из уравнений

$$\beta(\alpha)=C(1-\varepsilon)(\sigma+h\|\bar z\|),\quad \beta(\alpha)=C(\sigma+h\|\bar z\|),$$

то будет выполнено неравенство $0<\alpha_1(\delta)\leqslant\alpha_\delta\leqslant\alpha_2(\delta)$. Но тогда по лемме 2 имеем $\| z^{\alpha_2(\delta)}-z^{\alpha_\delta}\|\leqslant\| z^{\alpha_2(\delta)}-z^{\alpha_1(\delta)}\|$, и, значит, с учетом леммы 1

$$\| z^{\alpha_\delta}-\bar z\|\leqslant\| z^{\alpha_\delta}-z^{\alpha_2(\delta)}\|+\| z^{\alpha_2(\delta)}-\bar z\|\leqslant\| z^{\alpha_2(\delta)}-z^{\alpha_1(\delta)}\|+$$

$$+\| z^{\alpha_2(\delta)}-\bar z\|\leqslant 2\| z^{\alpha_2(\delta)}-\bar z\|+\| z^{\alpha_1(\delta)}-\bar z\|=O(\sigma+h).$$

Теорема доказана.

Отметим, что из результатов, которые приведены в § 1.13, 2.6, следует оптимальность по порядку алгоритма о.п.н. и в случае некоторых несовместных систем (1), а именно при условии единственности псевдорешения.

В связи со сказанным уточним содержание основной задачи. Нас будет интересовать построение такого устойчивого метода приближенного нахождения нормального псевдорешения системы (1), который имеет оптимальный порядок точности независимо от того, совместна эта система или нет.

Оказывается, что такой метод может быть создан на основе алгоритма о.п.н. (о.м.н.) при специальном выборе функционала Ω. Этот метод [131], обладающий рядом оптимальных свойств, изучается дальше.

§ 2. Метод минимальной псевдообратной матрицы

Рассмотрим СЛАУ (1.1) и будем считать, что матрица $\bar{A}$ принадлежит некоторому замкнутому в пространстве $\mathfrak{A}$ множеству $\mathfrak{A}_0$. В частности, возможен случай, когда $\mathfrak{A}_0=\mathfrak{A}$. Включение $\bar{A}\in\mathfrak{A}_0$ можно рассматривать как априорную информацию (если таковая имеется) о специальной структуре матрицы $\bar{A}$.

Используем приближенные данные (A_h, h) задачи (1.1) и введем при фиксированном h новое множество матриц

$$\mathfrak{A}_h\equiv\{A\in\mathfrak{A}_0\colon \|A-A_h\|\leqslant h\}.$$

В основе исследуемого в данном параграфе метода решения основной задачи — *метода минимальной псевдообратной матрицы* (м.п.м.) — лежит экстремальная задача: найти такую матрицу $\tilde{A}_h\in\mathfrak{A}_h$, для которой выполнено равенство

$$\|\tilde{A}_h^+\|_*=\inf\{\|A^+\|_*\colon A\in\mathfrak{A}_h\}\equiv\varkappa_h. \tag{1}$$

Будем называть любое решение $\tilde{A}_h$ задачи (1) *матрицей метода м.п.м.*, а соответствующую псевдообратную матрицу $\tilde{A}_h^+$ — *минимальной псевдообратной матрицей*, отвечающей данным (A_h, h).

Образуем с помощью любой минимальной псевдообратной матрицы $\tilde{A}_h^+$, отвечающей данным (A_h, h), элемент $z_\eta\equiv\tilde{A}_h^+u_\sigma$ $(\eta\equiv(h, \sigma))$ Этот элемент примем в качестве приближения по методу м.п.м. к нормальному псевдорешению $\bar{z}$ СЛАУ (1.1).

Перейдем теперь к обоснованию возможности реализации такого метода и к исследованию основных его свойств. Для этого понадобится

Лемма 1. *Пусть последовательность матриц $\{C_n\}\subset\mathfrak{A}$ сходится при $n\to\infty$ к $C_0\in\mathfrak{A}$, причем выполнено условие равномерной ограниченности норм*: $\|C_n^+\|_*\leqslant K=\text{const}$. *Тогда справедливо предельное соотношение $C_n^+\to C_0^+$ при $n\to\infty$.*

Доказательство. По условию равномерной ограниченности норм $\|C_n^+\|_*$ в конечномерном пространстве матриц $\mathfrak{A}^*$ множество $\{C_n^+\}$ будет компактно в $\mathfrak{A}^*$. Поэтому из последовательности $\{C_n^+\}$ можно извлечь сходящуюся подпоследовательность. Пусть $\{C_{n_l}^+\}$ — любая из них, т. е. $C_{n_l}^+\to D_0\in\mathfrak{A}^*$ при $l\to\infty$. Докажем, что предел D_0 совпадает с C_0^+.

Возьмем произвольный элемент $u\in\mathbf{R}^m$ и образуем последовательность нормальных псевдорешений $z_l=C_{n_l}^+u$ систем вида

$$C_{n_l}z=u. \tag{2}$$

Как известно (см., например, § 0.6), элемент z_l удовлетворяет уравнению

$$C_{n_l}^TC_{n_l}z_l=C_{n_l}^Tu. \tag{3}$$

Из сходимости последовательности $\{C_{n_l}^+\}$ к D_0 следует, что $z_l = C_{n_l}^+ u$ сходится при $l \to \infty$ к $z_0 \equiv D_0 u$. Это вместе с данной по условиям леммы сходимостью $C_{n_l} \to C_0$ позволяет получить из уравнения (3) путем предельного перехода при $l \to \infty$ следующее равенство:

$$C_0^T C_0 z_0 = C_0^T u.$$

Последнее уравнение означает, что z_0 — псевдорешение СЛАУ $C_0 z = u$.

Установим теперь, что z_0 — нормальное псевдорешение этой системы. Для этого достаточно показать, что $z_0 \perp \operatorname{Ker} C_0$, т. е. найдется элемент $v_0 \in \mathbf{R}^m$, для которого выполнено равенство $z_0 = C_0^T v_0$ (см. теорему 0.6.1). Для нормальных псевдорешений z_l систем вида (2) существуют элементы $v_l \in \mathbf{R}^m$ такие, что $z_l = C_{n_l}^T v_l$. Выбор подобных элементов v_l неоднозначен. В частности, удобно положить $v_l = (C_{n_l}^T)^+ z_l$. Тогда из сходимостей $C_{n_l}^+ \to D_0$, $z_l \to z_0$, а также из равенства $(C^T)^+ = (C^+)^T$ выводится предельное соотношение $v_l = (C_{n_l}^T)^+ z_l \to D_0^T z_0 \equiv v_0 \in \mathbf{R}^m$. Наконец, переходя к пределу в равенствах $C_{n_l}^T v_l = z_l$, получаем $C_0^T v_0 = z_0$. Это доказывает, что z_0 представляет собой нормальное псевдорешение системы $C_0 z = u$.

Итак, матрица D_0 позволяет по каждому $u \in \mathbf{R}^m$ вычислить нормальное псевдорешение $z_0 = D_0 u$ системы $C_0 z = u$. По определению псевдообратной матрицы и с использованием ее единственности можно сделать вывод, что $D_0 = C_0^+$. Таким образом, $C_{n_l}^+ \to C_0^+$ при $l \to \infty$. Из произвольности сходящейся последовательности $C_{n_l}^+$ вытекает, что C_0^+ — единственная предельная точка последовательности $\{C_n^+\}$.

Следствие 1. *Если для последовательности $\{C_n\}$, где $C_n \in \mathfrak{A}$, справедливы равномерные по n оценки $\|C_n\| \leqslant M_1$, $\|C_n^+\|_* \leqslant M_2$ ($M_{1,2} = \mathrm{const} > 0$), то множество $\{C_n\}$ компактно в $\mathfrak{A}$, а множество $\{C_n^+\}$ компактно в $\mathfrak{A}^*$.*

Рассмотрим проблему разрешимости задачи (1).

Теорема 1. *Решение задачи* (1) *существует для любых данных* (A_h, h), *где* $A_h \in \mathfrak{A}$, $h \geqslant 0$.

Доказательство. Пусть $\{A_N\}$ — произвольная минимизирующая последовательность матриц задачи (1): $A_N \in \mathfrak{A}_h$, $\|A_N^+\|_* \to \varkappa_h$ при $N \to \infty$. Тогда из неравенства $\|A_N - A_h\| \leqslant h$ вытекает оценка $\|A_N\| \leqslant \|A_h\| + h \equiv M_1$. С другой стороны, из сходимости $\|A_N^+\|_* \to \varkappa_h$ при $N \to \infty$ следует ограниченность норм $\|A_N^+\|_* \leqslant M_2$. Согласно следствию 1, множество $\{A_N\} \subset \mathfrak{A}_h$ компактно в $\mathfrak{A}$, а так как множество $\mathfrak{A}_h$ замкнуто в $\mathfrak{A}$, то существует подпоследовательность $\{A_{N_l}\}$, сходящаяся при $l \to \infty$ к матрице A_0, принадлежащей множеству $\mathfrak{A}_h$. С другой стороны, по лемме 1 имеет место сходимость $A_{N_l}^+ \to A_0^+$. Из нее и из определения минимизирующей для задачи (1) последовательности $\{A_{N_l}\}$ вытекает соотношение

$$\lim_{l \to \infty} \|A_{N_l}^+\|_* = \|A_0^+\|_* = \varkappa_h = \inf\{\|A^+\|_* \colon A \in \mathfrak{A}_h\}.$$

Таким образом, матрица A_0 есть решение задачи (1).

Теорема 2. *Пусть $\mathfrak{A}_0$ — подпространство в $\mathfrak{A}$. Если выполнено условие $\|A_h\| > h$, то все решения $\tilde{A}_h$ задачи* (1) *удовлетворяют равенству $\|\tilde{A}_h - A_h\| = h$.*

Доказательство. Предположим, что теорема неверна, и для некоторого решения $\tilde{A}_h$ задачи (1) имеет место строгое неравенство $\|\tilde{A}_h - A_h\| < h$. Введем функцию $f(\lambda) \equiv \|\lambda\tilde{A}_h - A_h\|$ и рассмотрим ее при $\lambda \geqslant 1$. Тогда по предположению оказывается, что $f(1) < h$. С другой стороны, по условию теоремы $\|A_h\| > h$, откуда ясно, что нулевая матрица не является элементом множества $\mathfrak{A}_h$. Поэтому и рассматриваемое решение $\tilde{A}_h$ задачи (1) не может быть нулевым. Следовательно,

$$f(\lambda) = \lambda \, \|\tilde{A}_h - (\lambda^{-1}) A_h\| \to +\infty,$$

если $\lambda \to +\infty$. Это вместе с непрерывностью функции $f(\lambda)$ при $\lambda \geqslant 1$ доказывает существование числа $\lambda_h > 1$ — корня уравнения $f(\lambda) = h$. Поскольку по условию $\mathfrak{A}_0$ — подпространство в $\mathfrak{A}$, то $\lambda_h \tilde{A}_h \in \mathfrak{A}_0$, $\|\lambda_h \tilde{A}_h - A_h\| = h$, т. е. $\lambda_h \tilde{A}_h \in \mathfrak{A}_h$. Кроме того, очевидно соотношение

$$\|(\lambda_h \tilde{A}_h)^+\|_* = \lambda_h^{-1} \|\tilde{A}_h^+\|_* < \|\tilde{A}_h^+\|_*.$$

Последнее неравенство, однако, противоречит тому, что $\tilde{A}_h^+$ — решение задачи (1). Это и доказывает теорему.

Подчеркнем, что условие теоремы $\|A_h\| > h$ выполнено при достаточно малых h, так как по предположению постановки задачи $\bar{A} \neq \mathbf{0}$.

Следствие 2. *При выполнении условий теоремы 2 задача (1) эквивалентна экстремальной задаче: найти такую матрицу $\tilde{A}_h \in \mathfrak{A}_0$, что*

$$\|\tilde{A}_h^+\|_* = \inf\{\|A^+\|_* : A \in \mathfrak{A}_0, \ \|A - A_h\| = h\}. \tag{4}$$

Заметим, что доказательство теоремы 2 верно и в том случае, если $\mathfrak{A}_0$ — конус в $\mathfrak{A}$.

Приведем примеры нахождения матрицы м.п.м. в простейшем случае.

Пример 1. Пусть

$$\bar{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_h = \begin{bmatrix} 1 & 0 \\ 0 & H \end{bmatrix}, \quad 0 < H < h;$$

$$\mathfrak{A}_0 \equiv \left\{ \begin{bmatrix} 1 & 0 \\ 0 & x \end{bmatrix} : x \in \mathbf{R} \right\} \subset \mathfrak{A} \equiv \mathfrak{A}_2.$$

Для матриц из класса $\mathfrak{A}_0$ легко находятся псевдообратные:

$$\begin{bmatrix} 1 & 0 \\ 0 & x \end{bmatrix}^+ = \begin{bmatrix} 1 & 0 \\ 0 & \theta(x) \end{bmatrix} = \operatorname{diag}[1, \theta(x)].$$

Здесь $\theta(x) = \{x^{-1}$ при $x \neq 0$; 0 при $x = 0\}$. Функция $\theta(x)$ в дальнейшем будет иметь большое значение для вычисления матриц м.п.м.

Задача (1) о нахождении минимальной псевдообратной матрицы $\tilde{A}_h = \operatorname{diag}[1, \theta(\tilde{x})]$ сводится к нахождению такого числа $\tilde{x} \in \mathrm{R}$, что

$$1 + \theta(\tilde{x}^2) = \inf\{1 + \theta(x^2) : (x - H)^2 \leqslant h^2\}.$$

На рис. 20 представлен график функции $1 + \theta(x^2)$. Легко видеть, что $1 + \theta(\tilde{x}^2) = \min\{1 + (H - h)^{-2}, 1, 1 + (H + h)^{-2}\} = 1$, т. е. $\tilde{x} = 0$.

Таким образом, $\tilde{A}_h = \operatorname{diag}(1,\ 0) = \bar{A}$. Отметим, что в данном примере отсутствует равенство $\|\tilde{A}_h - A_h\| = h$, и это связано со структурой множества $\mathfrak{A}_0$.

Пример 2. В примере 1 изменим множество $\mathfrak{A}_0$, положив $\mathfrak{A}_0 = \{\operatorname{diag}(x,\ y)\colon\ x,\ y \in \mathrm{R}\}$. Тогда для $A \in \mathfrak{A}_0$ можно найти $A^+ = \operatorname{diag}[\theta(x),\ \theta(y)]$. Предположим, что $\|A_h\|^2 = 1 + H^2 > h^2$ (например, рассмотрим h, подчиненные неравенству $0 \leqslant h < 1$). Тогда из следствия 2, учитывая тот факт, что $\mathfrak{A}_0$ — подпространство в $\mathfrak{A}$, можно найти минимальную псевдообратную матрицу $\tilde{A}_h^+ = \operatorname{diag}[\theta(\tilde{x}),\ \theta(\tilde{y})]$, решая задачу: найти числа $\tilde{x}$, $\tilde{y}$, для которых

$$\theta(\tilde{x}^2) + \theta(\tilde{y}^2) = \inf\{\theta(x^2) + \theta(y^2)\colon (x-1)^2 + (y-H)^2 = h^2\}.$$

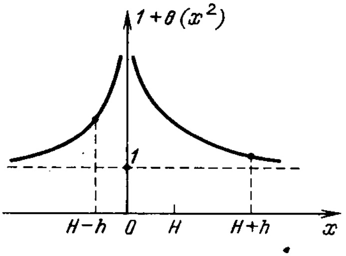

Рис. 20. Иллюстрация к примеру 1

Решение задач такого вида будет подробно рассмотрено в § 3 (см. пример 3.1). А пока ограничимся тем, что выпишем решение

$$\tilde{x} = 1 + (h^2 - H^2)^{1/2}, \quad \tilde{y} = 0, \quad 0 \leqslant H \leqslant h < 1/2.$$

Теорема 3. *Если $\tilde{A}_h$ — матрица метода м.п.м., то $\tilde{A}_h^+$ сходится при $h \to 0$ к $\bar{A}^+$.*

Доказательство. Заметим, что из включения $\bar{A} \in \mathfrak{A}_0$ и из (1) следует оценка

$$\|\tilde{A}_h^+\|_* \leqslant \|\bar{A}^+\|_*. \tag{5}$$

Поскольку $A_h \to \bar{A}$ при $h \to 0$, отсюда по лемме 1 и получим доказываемое утверждение.

Результат теоремы 3 позволяет обосновать сходимость приближений $z_\eta = \tilde{A}_h^+ u_\sigma$ метода м.п.м. к нормальному псевдорешению $\bar{z}$ системы (1.1) при $\eta \to 0$. Действительно, очевидна оценка

$$\begin{aligned}\|z_\eta - \bar{z}\| &= \|\tilde{A}_h^+ u_\sigma - \bar{A}^+ \bar{u}\| \leqslant \|\tilde{A}_h^+ u_\sigma - \tilde{A}_h^+ \bar{u}\| + \\ &+ \|\tilde{A}_h^+ \bar{u} - \bar{A}^+ \bar{u}\| \leqslant \|\tilde{A}_h^+\|_* \cdot \|u_\sigma - \bar{u}\| + \|\tilde{A}_h^+ - \bar{A}^+\|_* \cdot \|\bar{u}\| \leqslant \\ &\leqslant \|\bar{A}^+\|_* \sigma + \|\tilde{A}_h^+ - \bar{A}^+\|_* \cdot \|\bar{u}\|, \end{aligned} \tag{6}$$

для получения которой использовано неравенство (5). Переходя в оценке (6) к пределу при $\eta = (h,\ \sigma) \to 0$, получаем сходимость $z_\eta \to \bar{z}$. Ниже будет получена оценка точности приближений метода м.п.м. А пока исследуем некоторые его свойства.

Прежде всего отметим, что матрица метода м.п.м. $\tilde{A}_h$ подчиняется неравенству

$$\|\tilde{A}_h - \bar{A}\| \leqslant \|\tilde{A}_h - A_h\| + \|A_h - \bar{A}\| \leqslant 2h, \tag{7}$$

которое используется в дальнейшем. Далее, одним из важнейших свойств матрицы метода м.п.м. оказывается совпадение ее ранга с рангом неизвестной точной матрицы $\bar{A}$ при достаточно малых h.

Теорема 4. *Пусть* $\tilde{A}_h$ — *любое решение задачи* (1) *при* $0 \leqslant h < h_0(\bar{A}) \equiv \| \bar{A}^+ \|_*^{-1}/2$. *Тогда* $\operatorname{Rg} \tilde{A}_h = \operatorname{Rg} \bar{A}$.

Прежде чем приступить к доказательству, введем некоторые величины, используемые для исследования. Обозначим $M \equiv \min(m,n)$ и рассмотрим сингулярные разложения матриц $\bar{A}, A_h$: $\bar{A} = \bar{U}\bar{D}\bar{V}^T$, $A_h = U_h D_h V_h^T$. Здесь $\bar{U}$, $U_h \in \mathfrak{A}_m$, $\bar{V}$, $V_h \in \mathfrak{A}_n$ — ортогональные матрицы, а $\bar{D} = \operatorname{diag}(\bar{\rho}_1, \ldots, \bar{\rho}_M) \in \mathfrak{A}$, $D_h = \operatorname{diag}(\rho_1^h, \ldots, \rho_M^h) \in \mathfrak{A}$ — прямоугольные диагональные матрицы, содержащие упорядоченные по невозрастанию сингулярные числа $\bar{\rho}_k$, ρ_k^h $(k = 1, \ldots, M)$ матриц $\bar{A}$, A_h соответственно: $\bar{\rho}_1 \geqslant \ldots \geqslant \bar{\rho}_M \geqslant 0$, $\rho_1^h \geqslant \ldots \geqslant \rho_M^h \geqslant 0$.

Как известно (см. [46, 143] и § 0.6), для сингулярных чисел $\rho_1 \geqslant \ldots \geqslant \rho_M$ произвольной матрицы $A \in \mathfrak{A}$ выполнены соотношения

$$\sum_{k=1}^{M} (\rho_k - \bar{\rho}_k)^2 \leqslant \| A - A_h \|^2, \tag{8}$$

$$\| A^+ \|_*^2 = \sum_{k=1}^{M} \theta(\rho_k^2) = \sum_{k=1}^{r(A)} \rho_k^{-2}. \tag{9}$$

При этом $\rho_k = 0$ для $r(A) < k \leqslant M$, где $r(A) \equiv \operatorname{Rg} A$ — ранг матрицы A.

Доказательство теоремы 4. Пусть $\tilde{\rho}_1, \ldots, \tilde{\rho}_M$ — сингулярные числа матрицы метода м.п.м. $\tilde{A}_h$. Предположим, что $t \equiv \operatorname{Rg} \tilde{A}_h$, $r \equiv \operatorname{Rg} \bar{A}$, причем $t < r$. Тогда из неравенств (7), (8) получим

$$\bar{\rho}_r^2 \leqslant \sum_{k=1}^{t} (\tilde{\rho}_k - \bar{\rho}_k)^2 + \sum_{k=t+1}^{r} \bar{\rho}_k^2 = \sum_{k=1}^{M} (\tilde{\rho}_k - \bar{\rho}_k) \leqslant \| \tilde{A}_h - \bar{A} \|^2 \leqslant 4h^2,$$

а из (9) получаем

$$\bar{\rho}_r^{-2} \leqslant \| \bar{A}^+ \|_*^2.$$

Из этих двух оценок для $\bar{\rho}_r$ следует, что $h^2 \geqslant \| \bar{A}^+ \|^{-2}/4$, что противоречит условию теоремы. Поэтому должно выполняться неравенство $t \geqslant r$. Если теперь допустить, что оно строгое $t > r$, то снова получим из (7)—(9) с учетом (5), что

$$\tilde{\rho}_{r+1}^2 = (\tilde{\rho}_{r+1} - \bar{\rho}_{r+1})^2 \leqslant \sum_{k=1}^{M} (\tilde{\rho}_k - \bar{\rho}_k)^2 \leqslant 4h^2,$$

$$\tilde{\rho}_{r+1}^{-2} \leqslant \sum_{k=1}^{t} \tilde{\rho}_k^{-2} = \| \tilde{A}_h^+ \|_*^2 \leqslant \| \bar{A}^+ \|_*^2.$$

Отсюда снова следует, что $h^2 \geqslant h_0^2(\bar{A})$. Это противоречие с требованием на число h доказывает, что $t = r$.

Следствие 3. *Сингулярные числа* $\tilde{\rho}_k$ *матрицы* $\tilde{A}_h$ *равны нулю при* $k > r$ *и отличны от нуля при* $1 \leqslant k \leqslant r$, *если* $0 \leqslant h < h_0(\bar{A})$.

Следствие 4. *Рассуждая, как при доказательстве теоремы* 4 (*случай* $t < r$), *можно установить, что* $\operatorname{Rg} A_h \geqslant \operatorname{Rg} \bar{A}$.

По аналогии с доказательством теоремы 4 устанавливается

Следствие 5. *Если для параметрического семейства матриц* $\{\breve{A}_h\}$, *где* $h \geqslant 0$, *выполнены неравенства* $\| \breve{A}_h - \bar{A} \| \leqslant qh$ $(q = \text{const} \geqslant 1)$, $\| \breve{A}_h^+ \|_* \leqslant \| \bar{A}^+ \|_*$, *то при выполнении условия* $0 \leqslant h < \| \bar{A}^+ \|_*^{-1}/q$ *справедливо равенство* $\operatorname{Rg} \breve{A}_h = \operatorname{Rg} \bar{A}$.

Из теоремы 4 ясно, что $\operatorname{Rg}\tilde{A}_h^+ = \operatorname{Rg}\bar{A}^+$ при $0 \leqslant h < h_0\ (\bar{A})$. Это свойство отличает метод м.п.м. от многих других алгоритмов устойчивого решения СЛАУ, использующих аппроксимацию псевдообратной матрицы. Так, например, в методе из работ [28, 55, 164] применяется следующая аппроксимация для $\bar{A}^+$:

$$B_h = (hI + A_h^T A_h)^{-1} A_h^T A_h A_h^T (hE + A_h A_h^T)^{-1}.$$

Очевидно, что матрица B_h может при сколь угодно малых h иметь ранг, отличный от $\operatorname{Rg}\bar{A}$. Чтобы убедиться в этом, достаточно рассмотреть матрицы A_h, $\bar{A}$ из примера 1.

Используя выводы теоремы 4, можно дать оценку скорости сходимости $\tilde{A}_h^+ \to \bar{A}^+$ при $h \to 0$, установленной в теореме 3. При этом можно воспользоваться следующим результатом [242].

Лемма 2. *Пусть $A, B \in \mathfrak{A}$, $\operatorname{Rg} A = \operatorname{Rg} B$. Тогда*

$$\| A^+ - B^+ \|_* \leqslant L \| A - B \| \cdot \| A^+ \|_* \cdot \| B^+ \|_*, \tag{10}$$

где константа L определяется следующим образом: $L = 1$ при $\operatorname{Rg} A = M$, $L = \sqrt{2}$ при $\operatorname{Rg} A < M$.

Поскольку, согласно теореме 4, $\operatorname{Rg}\tilde{A}_h = \operatorname{Rg}\bar{A}$ при $0 \leqslant h < h_0(\bar{A})$, то из леммы 2 и из (5), (7) следует

Теорема 5. *При условии $0 \leqslant h < h_0(\bar{A})$ справедлива оценка*

$$\| \tilde{A}_h^+ - \bar{A}^+ \|_* \leqslant \sqrt{2} \| \tilde{A}_h - \bar{A} \| \cdot \| \tilde{A}_h^+ \|_* \cdot \| \bar{A}^+ \|_* \leqslant 2\sqrt{2} h \| \bar{A}^+ \|_*^2.$$

Эта оценка, как нетрудно убедиться на простейших примерах, является точной по порядку величин h, $\| \bar{A}^+ \|_*$. Константа $2\sqrt{2}$, однако, точной не является и может быть улучшена при $h \to 0$. Величина этой константы оказывается существенной при исследовании оптимальных свойств точности метода м.п.м. Оценку (10) уточняет

Лемма 3. *Пусть $A, B \in \mathfrak{A}$, $\operatorname{Rg} A = \operatorname{Rg} B$, $\| A - B \| \cdot \| A^+ \|_* < 1$. Тогда*

$$\| A^+ - B^+ \|_* \leqslant \| A - B \| \cdot \| A^+ \|_*^2 / (1 - \| A - B \| \cdot \| A^+ \|_*)^3.$$

Доказательство. Пусть матрица A имеет сингулярное разложение $A = UPV^T$, где $P = \operatorname{diag}(p_1, \ldots, p_t, 0, \ldots, 0) \in \mathfrak{A}$, $t \equiv \operatorname{Rg} A$. Тогда в силу ортогональности матриц U и V, а также вследствие равенства $A^+ = VP^+U^T$ получим

$$\begin{gathered} \| A - B \| = \| UPV^T - B \| = \| P - U^T BV \| = \| P - Q \|, \\ Q \equiv U^T BV, \quad \| A^+ - B^+ \|_* = \| VP^+U^T - B^+ \|_* = \\ = \| P^+ - V^T B^+ U \|_* = \| P^+ - (U^T BV)^+ \|_* = \| P^+ - Q^+ \|_*, \\ \| A^+ \|_* = \| P^+ \|_*. \end{gathered} \tag{11}$$

Поскольку $\operatorname{Rg} P = t = \operatorname{Rg} B = \operatorname{Rg} Q$, то матрицы P, Q можно представить в следующем блочном виде:

$$\begin{aligned} &P = [W \vdots 0], \quad W \in \mathfrak{A}_{m \times t}, \quad 0 \in \mathfrak{A}_{m \times (n-t)}, \quad \operatorname{Rg} W = t, \\ &Q = [R \vdots S], \quad R \in \mathfrak{A}_{m \times t}, \quad S \in \mathfrak{A}_{m \times (n-t)}, \quad \operatorname{Rg} R = \operatorname{Rg} Q = t, \end{aligned}$$

причем вследствие линейной зависимости столбцов матрицы S от столбцов матрицы R (теорема о базисном миноре) будет выполняться равенство $S=R^+RS$. Как показано в [7], при этих условиях справедливы равенства

$$P^+=\begin{bmatrix} W^+ \\ 0 \end{bmatrix}, \quad Q^+=\begin{bmatrix} R^+-R^+SJ \\ J \end{bmatrix},$$

где

$$J\equiv KZ^TR^+\in\mathfrak{A}_{(n-t)\times m}, \quad K\equiv(E+Z^TZ)^{-1},$$
$$Z\equiv R^+S\in\mathfrak{A}_{n\times(n-t)}, \quad E\equiv\operatorname{diag}(1, \ldots, 1)\in\mathfrak{A}_{(n-t)\times(n-t)}.$$

Поэтому

$$\|P^+-Q^+\|_*^2=\|W^+-R^++R^+SJ\|^2+$$
$$+\|J\|^2\leqslant(\|W^+-R^+\|+\|R^+SJ\|)^2+\|J\|^2.$$

Оценим правую часть этого неравенства. Прежде всего, справедлива выкладка

$$\|J\|^2=\|KZ^TR^+\|^2\leqslant\|(E+Z^TZ)^{-1}Z^T\|^2\cdot\|R^+\|^2=$$
$$=\|R^+\|^2\sum_{i=1}^{r(Z)}\frac{\mu_i^2}{(1+\mu_i^2)^2}\leqslant\|R^+\|^2\sum_{i=1}^{r(Z)}\mu_i^2=$$
$$=\|R^+\|^2\cdot\|Z\|^2=\|R^+\|^2\cdot\|R^+S\|^2\leqslant\|R^+\|^4\cdot\|S\|^2.$$

Здесь μ_i $(i=1, \ldots, r(Z))$ — ненулевые сингулярные числа матрицы Z, $r(Z)\equiv\operatorname{Rg}Z$.

Используя полученную оценку, имеем далее

$$\|R^+SJ\|\leqslant\|R^+\|\cdot\|S\|\cdot\|J\|\leqslant\|R^+\|^3\cdot\|S\|^2.$$

Таким образом,

$$\|P^+-Q^+\|_*^2\leqslant(\|W^+-R^+\|+\|R^+\|^3\cdot\|S\|^2)^2+\|R^+\|^4\cdot\|S\|^2. \quad (12)$$

Оценим теперь величины $\|W^+-R^+\|$, $\|R^+\|$. По лемме 2 для W, R как для матриц полного ранга t выполнено неравенство

$$\|W^+-R^+\|\leqslant\|W^+\|\cdot\|R^+\|\cdot\|W-R\|. \quad (13)$$

Из условия леммы и соотношений (11) получим

$$\|W-R\|^2\cdot\|W^+\|^2\leqslant(\|W-R\|^2+\|S\|^2)\cdot\|W^+\|^2=$$
$$=\|P-Q\|^2\cdot\|P^+\|_*^2=\|A-B\|^2\cdot\|A^+\|_*^2<1,$$

вследствие чего из (13) вытекает оценка

$$\|R^+\|\leqslant\|W^+\|/(1-\|W-R\|\cdot\|W^+\|). \quad (14)$$

Обозначим для краткости

$$\|W-R\|\equiv\varepsilon, \quad \|S\|\equiv\omega, \quad \|A-B\|^2=\|P-Q\|^2=\varepsilon^2+\omega^2\equiv H^2.$$

Тогда из (11), (12) с использованием (13), (14) получается цепочка неравенств:

$$\|A^+-B^+\|_*=\|P^+-Q^+\|_*\leqslant$$
$$\leqslant[\|W^+\|^2\varepsilon(1-\varepsilon\|W^+\|)^{-1}+\|W^+\|^3\omega^2(1-\varepsilon\|W^+\|)^{-3}]^2+$$
$$+\|W^+\|^4\omega^2(1-\varepsilon\|W^+\|)^{-4}=$$

$$=\frac{\| W^+ \|^4}{(1-\varepsilon \| W^+ \|)^4}\left[\varepsilon^2(1-\varepsilon \| W^+ \|)^2+\omega^2+2\varepsilon\omega^2 \| W^+ \| +\frac{\omega^4 \| W^+ \|^2}{(1-\varepsilon \| W^+ \|)^2}\right]\leqslant$$

$$\leqslant\frac{\| W^+ \|^4(\varepsilon^2+\omega^2)}{(1-\varepsilon \| W^+ \|)^4}\left[1+\frac{\varepsilon^4 \| W^+ \|^2}{\varepsilon^2+\omega^2}+\frac{2\varepsilon\omega^2}{\varepsilon^2+\omega^2} \| W^+ \| +\frac{\omega^4}{\varepsilon^2+\omega^2}\,\frac{\| W^+ \|^2}{(1-\varepsilon \| W^+ \|)^2}\right]\leqslant$$

$$\leqslant\frac{\| W^+ \|^4 H^2}{(1-H \| W^+ \|)^4}\left[1+2\varepsilon \| W^+ \| +\frac{\varepsilon^4+\omega^4}{\varepsilon^2+\omega^2}\,\frac{\| W^+ \|^2}{(1-H \| W^+ \|)^2}\right]\leqslant$$

$$\leqslant\frac{H^2 \| W^+ \|^4}{(1-H \| W^+ \|)^4}\left[1+\frac{2H \| W^+ \|}{1-H \| W^+ \|}+\frac{H^2 \| W^+ \|^2}{(1-H \| W^+ \|)^2}\right]=$$

$$=\frac{H^2 \| W^+ \|^4}{(1-H \| W^+ \|)^4}\left[1+\frac{H \| W^+ \|}{1-H \| W^+ \|}\right]^2=\frac{H^2 \| W^+ \|^4}{(1-H \| W^+ \|)^6}=\frac{H^2 \| A^+ \|_*^4}{(1-H \| A^+ \|_*)^6},$$

где учтены также очевидные неравенства

$$1\leqslant(1-\varepsilon \| W^+ \|)^{-1}\leqslant(1-H \| W^+ \|)^{-1},\quad \varepsilon^2/(\varepsilon^2+\omega^2)\leqslant 1.$$

Лемма доказана.

Из леммы 3 и теоремы 4 вытекает

Теорема 6. *При условии* $0\leqslant h<h_0(\bar{A})$ *справедлива оценка*

$$\| \tilde{A}_h^+ - \bar{A}^+ \|_* \leqslant 2h \| \bar{A}^+ \|_*^2/(1-2h \| \bar{A}^+ \|_*)^3. \tag{15}$$

Ясно, что при $h\to 0$ главный член оценки (15) имеет вид

$$\| \tilde{A}_h^+ - \bar{A}^+ \|_* \asymp 2h \| \bar{A}^+ \|_*^2.$$

Константа 2 здесь асимптотически неулучшаема: найдутся такие матрицы $\bar{A}$, A_h, подчиненные условию $\| \bar{A}-A_h \| =h$, для которых оказывается выполненным равенство

$$\lim_{h\to+0} \| \tilde{A}_h^+ - \bar{A}^+ \|_*/(h \| \bar{A}^+ \|_*^2)=2.$$

Чтобы убедиться в этом, рассмотрим

Пример 3. Пусть $\bar{A}=\mathrm{diag}\,(1,\,0)\in\mathfrak{A}_2$ и множество $\mathfrak{A}_0$ определено следующим образом:

$$\mathfrak{A}_0=\left\{\begin{bmatrix}1 & x\\ x & x^2\end{bmatrix}: x\in\mathbf{R}\right\}.$$

Зададим также A_h:

$$A_h=\begin{bmatrix}1 & H\\ H & H^2\end{bmatrix},\quad 2H^2+H^4\equiv h^2,\quad 0\leqslant h<1,\quad H\geqslant 0.$$

Очевидно, что выполнено равенство $\| \bar{A}-A_h \|^2=h^2$. Для любой матрицы из множества $\mathfrak{A}_0$ можно вычислить псевдообратную:

$$A^+=\begin{bmatrix}1 & x\\ x & x^2\end{bmatrix}^+=\lim_{\alpha\to+0}(\alpha E+A^TA)^{-1}A^T=\frac{1}{(1+x^2)^2}\begin{bmatrix}1 & x\\ x & x^2\end{bmatrix},$$

причем

$$\| A^+ \|_*^2=\frac{1+2x^2+x^4}{(1+x^2)^4}=\frac{1}{(1+x^2)^2}.$$

Множество $\mathfrak{A}_h$ определяется неравенством $\| A-A_h \|^2=2(x-H)+(x^2-H^2)^2\leqslant h^2=2H^2+H^4$. Значит, $A\in\mathfrak{A}_h$, если выполнено неравенство $0\leqslant x\leqslant x(H)$, где $x(H)$ — единственный корень уравнения $x^3+2x(1-H^2)-4H=0$. Нетрудно найти, что $x(H)\asymp 2H$ при $H\to 0$. Тогда задача (1) о нахождении матрицы метода м.п.м. сводится к определению такого числа $\tilde{x}$, для которого

$$(1+\tilde{x}^2)^{-2}=\inf\{(1+x^2)^{-2}\colon 0\leqslant x\leqslant x(H)\}.$$

Поэтому $\tilde{x}=x(H)$, и далее с учетом равенства $\| \bar{A}^+ \|_*=1$ получаем

$$\| \tilde{A}_h^+ - \bar{A}^+ \|_*^2=\left(1-\frac{1}{(1+\tilde{x}^2)^2}\right)^2+\frac{2\tilde{x}^2}{(1+\tilde{x}^2)^4}+\frac{\tilde{x}^4}{(1+\tilde{x}^2)^4}=\frac{2\tilde{x}^2+\tilde{x}^4}{(1+\tilde{x}^2)^2}\ \| \bar{A}^+ \|_*^4,$$

откуда

$$\lim_{h\to+0}\frac{\| \tilde{A}_h^+ - \bar{A}^+ \|_*}{h\| \bar{A}^+ \|_*^2}=\lim_{H\to 0}\sqrt{\frac{2x^2(H)+x^4(H)}{(2H^2+H^4)\,[1+x^2(H)]^2}}=2.$$

На основании теоремы 6 можно получить оценку точности приближений z_η метода м.п.м. к нормальному псевдорешению $\bar{z}$. Для этого используем неравенство (6).

Теорема 7. *Справедлива оценка*

$$\| z_\eta-\bar{z} \| \leqslant \| \bar{A}^+ \|_* \sigma + \| \tilde{A}_h^+ - \bar{A}^+ \|_* \| \bar{u} \| \leqslant$$
$$\leqslant \| \bar{A}^+ \|_* \left[\sigma+2h \| \bar{A}^+ \|_* \| \bar{u} \| /(1-2h\| \bar{A}^+ \|_*)^3\right]. \qquad (16)$$

Из оценки (16) понятно, что приближение z_η имеет оптимальный порядок точности $O(\sigma+h)$ независимо от того, совместна система (1.1) или нет. Используя (16), можно сравнить оценку точности метода м.п.м. и оценку точности метода, исследованного в работах [28, 55]. Последняя дана в формуле (1.9).

В заключение заметим, что задача (1) метода м.п.м. по форме представляет собой вариант обобщенного метода невязки (см. § 2.5), в котором принято $\Omega=\| A^+ \|_*$.

Такой функционал $\Omega(A)$, однако, не обладает свойствами полунепрерывности и компактности в $\mathfrak{A}$, которые сформулированы в § 1.2. Поэтому теория обобщенного метода невязки из гл. 2 не переносится непосредственно на метод м.п.м.

Вместе с тем функционал $\Omega(A)$ обладает все же несколько другими специальными свойствами «ограниченной непрерывности» и «ограниченной компактности» в $\mathfrak{A}$, которые обоснованы в лемме 1 и следствии 1. Эти свойства в конечном итоге и обеспечивают сходимость метода м.п.м. (см. теорему 3).

§ 3. Нахождение матрицы метода м.п.м.

Центральным моментом реализации метода м.п.м. является решение задачи (2.1). При этом достаточно найти любое ее решение $\tilde{A}_h$ и с его помощью образовать приближенное решение $z_\eta=\tilde{A}_h^+ u_\sigma$ системы (1.1). В данном параграфе будет указан удобный для нахождения специальный вид решения задачи (2.1) для случая, когда в качестве априорного множества $\mathfrak{A}_0$ берется все пространство $\mathfrak{A}$: $\mathfrak{A}_0=\mathfrak{A}$.

Изучим прежде всего вспомогательную задачу математического программирования: при фиксированном h и при заданных сингулярных числах ρ_k^h ($k=1, \ldots, M$) матрицы A_h найти такие числа $\hat{\rho}_1, \ldots, \hat{\rho}_M$, которые удовлетворяют условию монотонности $\hat{\rho}_1 \geqslant \ldots \geqslant \hat{\rho}_M \geqslant 0$ и для которых выполнено равенство

$$\sum_{k=1}^{M} \theta(\hat{\rho}_k^2) = \\ = \inf \left\{ \sum_{k=1}^{M} \theta(\rho_k^2) \colon \rho_1 \geqslant \rho_2 \geqslant \ldots \geqslant \rho_M \geqslant 0, \quad \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2 = h^2 \right\}. \tag{1}$$

Наряду с этой задачей будем рассматривать также задачу: найти такие числа $\hat{\rho}_1, \ldots, \hat{\rho}_M$ ($\hat{\rho}_1 \geqslant \ldots \geqslant \hat{\rho}_M \geqslant 0$), для которых

$$\sum_{k=1}^{M} \theta(\hat{\rho}_k^2) = \inf \left\{ \sum_{k=1}^{M} \theta(\rho_k^2) \colon \rho_1 \geqslant \ldots \geqslant \rho_M \geqslant 0, \quad \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2 \leqslant h^2 \right\}. \tag{2}$$

Связь задач (1) и (2) устанавливает

Лемма 1. *Задачи* (1) *и* (2) *разрешимы и эквивалентны при выполнении условия* $\| A_h \| > h$.

Доказательство. Фиксируем данные (A_h, h) и используем матрицы U_h, V_h из сингулярного разложения матрицы A_h для того, чтобы определить множество матриц

$$\mathfrak{A}(A_h) = \{A \colon A \in \mathfrak{A}, \ A = U_h D V_h^T, \\ D = \operatorname{diag}(\rho_1, \ldots, \rho_M) \in \mathfrak{A}, \ \rho_1 \geqslant \ldots \geqslant \rho_M \geqslant 0\}.$$

Вследствие того, что для любой матрицы из множества $\mathfrak{A}(A_h)$ справедливо равенство

$$\| A - A_h \|^2 = \| U_h (D - D_h) V_h^T \|^2 = \| D - D_h \|^2 = \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2,$$

и с учетом равенства (2.9) можно сделать вывод, что задачи (1), (2) могут быть записаны в эквивалентной форме:

$$\inf \{ \| A^+ \|_*^2 \colon A \in \mathfrak{A}(A_h), \quad \| A - A_h \|^2 = h^2 \}, \tag{1'}$$

$$\inf \{ \| A^+ \|_*^2 \colon A \in \mathfrak{A}(A_h), \quad \| A - A_h \|^2 \leqslant h^2 \}. \tag{2'}$$

Задачи (1′), (2′) по своей форме соответствуют задачам (2.4), (2.1), если в последних рассматривать фиксированное h и положить $\mathfrak{A}_0 = \mathfrak{A}(A_h)$.

Поэтому существование решений задач (1′), (2′) и их эквивалентность при условии

$$\| A_h \|^2 = \sum_{k=1}^{M} (\rho_k^h)^2 > h^2 \tag{3}$$

получаются из теорем 2.1, 2.2.

Теорема 1. *Предположим, что при выполнении условия* (3) *числа* $\hat{\rho}_1, \ldots, \hat{\rho}_M$ *представляют собой какое-либо решение задачи* (1). *Тогда* *матрица* $\hat{A}_h \equiv U_h \hat{D}_h V_h^T$, *где* $\hat{D}_h \equiv \mathrm{diag}(\hat{\rho}_1, \ldots, \hat{\rho}_M) \in \mathfrak{A}$, *есть решение экстремальной задачи* (2.1) *метода м.п.м. с* $\mathfrak{A}_0 = \mathfrak{A}$.

Таким образом, одна из матриц метода м.п.м. определяется матрицами U_h, V_h, которые входят в сингулярное разложение заданной матрицы A_h, а также диагональной матрицей $\hat{D}_h$, содержащей решение задачи (1).

Доказательство. Рассмотрим произвольную матрицу $A \in \mathfrak{A}$, имеющую сингулярное разложение $A = UDV^T$ $(D = \mathrm{diag}(\rho_1, \ldots, \rho_M))$. Тогда вследствие формулы (2.8) справедливо соотношение

$$\| D - D_h \|^2 = \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2 \leqslant \| A - A_h \|^2 \quad \forall A \in \mathfrak{A}. \tag{4}$$

Введем теперь, как в § 2, множество $\mathfrak{A}_h = \{A \in \mathfrak{A}\colon \| A - A_h \| \leqslant h\}$, а также новое множество $\mathfrak{B}_h$, состоящее из матриц $A = UDV^T$, которые удовлетворяют ограничению $\| D - D_h \| \leqslant h$:

$$\mathfrak{B}_h = \{A \in \mathfrak{A}\colon A = UDV^T,\ \ \| D - D_h \| \leqslant h\}.$$

Поскольку, согласно (4), неравенство $\| A - A_h \| \leqslant h$ влечет за собой выполнение неравенства $\| D - D_h \| \leqslant h$, то $\mathfrak{A}_h \subset \mathfrak{B}_h$. Из этого вложения вытекает соотношение

$$\begin{aligned}
&\inf\{\| A^+ \|_*^2\colon A \in \mathfrak{A}_h\} \geqslant \\
&\geqslant \inf\{\| A^+ \|_*^2\colon A \in \mathfrak{B}_h\} = \inf\{\| D^+ \|_*^2\colon \| D - D_h \| \leqslant h\} = \\
&= \inf\left\{ \sum_{k=1}^{M} \theta(\rho_k^2)\colon \rho_1 \geqslant \ldots \geqslant \rho_M \geqslant 0,\ \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2 \leqslant h^2 \right\},
\end{aligned} \tag{5}$$

в котором использован вариант равенства (2.9):

$$\| A^+ \|_*^2 = \| D^+ \|_*^2 = \sum_{k=1}^{M} \theta(\rho_k^2),$$

а также соотношение (4). В формуле (5) по лемме 1 последняя точная нижняя грань достигается на числах $\hat{\rho}_1, \ldots, \hat{\rho}_M$ — решении задачи (1), так что из (5) следует неравенство

$$\| A^+ \|_*^2 \geqslant \inf\{\| A^+ \|_*^2\colon A \in \mathfrak{A}_h\} \geqslant \sum_{k=1}^{M} \theta(\hat{\rho}_k^2) = \| \hat{D}_h^+ \|_*^2 \quad \forall A \in \mathfrak{A}_h. \tag{6}$$

Введем теперь матрицу $\hat{A}_h \equiv U_h \hat{D}_h V_h^T$. Вследствие (1) получим

$$\| \hat{A}_h - A_h \|^2 = \| U_h \hat{D}_h V_h^T - U_h D_h V_h^T \|^2 = \| \hat{D}_h - D_h \|^2 = \sum_{k=1}^{M} (\hat{\rho}_k - \rho_k^h)^2 = h^2$$

и, значит, $\hat{A}_h \in \mathfrak{A}_h$. Но тогда, согласно неравенству (6),

$$\| \hat{A}_h^+ \|_*^2 = \| \hat{D}_h^+ \|_*^2 \leqslant \| A^+ \|_*^2 \quad \forall A \in \mathfrak{A}_h.$$

Таким образом, матрица $\hat{A}_h$ есть решение задачи (2.1). Теорема доказана.

Теорема 1 дает конструктивный способ определения матрицы метода м.п.м. Для его реализации необходимо прежде всего найти сингулярное разложение заданной матрицы A_h, а затем с его помощью, находя какое-нибудь решение $(\hat{\rho}_1, ..., \hat{\rho}_M)$ задачи (1), вычислить матрицу метода м.п.м. $\hat{A}_h$. Алгоритмические вопросы, связанные с нахождением сингулярного разложения, будут подробно обсуждены ниже. Сейчас же нас будет интересовать возможность получить в явной форме решение задачи (1). Изложим два возможных подхода к решению этой задачи.

Первый подход основан на использовании функции Лангранжа, связанной с задачей (1). Определим ее в следующем виде:

$$M^\lambda(\rho_1, ..., \rho_M) \equiv \lambda \sum_{k=1}^{M} \theta(\rho_k^2) + \sum_{k=1}^{M} (\rho_k - \rho_k^h)^2, \quad \lambda \geqslant 0, \quad \rho_1, ..., \rho_M \geqslant 0. \tag{7}$$

Будем рассматривать экстремальную задачу: найти при фиксированном $\lambda \geqslant 0$ такие числа $\rho_1(\lambda), ..., \rho_M(\lambda) \geqslant 0$, для которых

$$M^\lambda[\rho_1(\lambda), ..., \rho_M(\lambda)] = \inf\{M^\lambda[\rho_1, ..., \rho_M]: \rho_1, ..., \rho_M \geqslant 0\}. \tag{8}$$

Лемма 2. *Задача* (8) *разрешима при всяком* $\lambda \geqslant 0$. *Все ее решения определяются формулами*

$$\rho_k(\lambda) = \{\rho_k^h x_k(\lambda): 0 < \lambda \leqslant \lambda_k; \quad 0: \lambda \geqslant \lambda_k\}, \tag{9}$$

$$\rho_k(0) = \rho_k^h, \quad k = 1, ..., M, \tag{9'}$$

где $\lambda_k \equiv 27(\rho_k^h)^4/16$, $x_k(\lambda)$ — *решение уравнения*

$$x^4 - x^3 = \lambda(\rho_k^h)^{-4}, \tag{10}$$

лежащее на отрезке [1, 3/2]. *Величины* $\rho_k(\lambda)$, *определяемые согласно формулам* (9), (9′), *обладают свойством монотонности*: $\rho_1(\lambda) \geqslant ... \geqslant \rho_M(\lambda) \geqslant 0$.

Для доказательства этой и некоторых последующих лемм понадобятся некоторые свойства положительного решения параметрического уравнения

$$x^4 - x^3 = \lambda\rho^{-4}, \tag{10'}$$

где $\rho > 0$ — фиксированное число, а $\lambda \geqslant 0$ — параметр. Элементарное исследование этого уравнения показывает, что для каждого $\lambda \geqslant 0$ оно имеет единственное положительное решение $x(\lambda)$, причем выполнено неравенство $x(\lambda) \geqslant 1$. С помощью стандартных методов исследования функций нетрудно установить свойства функции $x(\lambda)$. Перечислим их.

1. Функция $x(\lambda)$ непрерывна и дифференцируема при $\lambda \geqslant 0$ любое число раз. В частности,

$$x'(\lambda) = \rho^{-4}/(4x^3 - 3x^2), \quad x''(\lambda) = -6\rho^{-8}(2x-1)/[x^5(4x-3)^3].$$

2. Функция $x(\lambda)$ монотонно возрастает при $\lambda \geqslant 0$, причем $x(0) = 1$, $x(+\infty) = +\infty$.

3. Функция $x(\lambda)$ выпукла вверх при $\lambda \geqslant 0$.

4. Величина $x \geqslant 1$, определяемая из рассматриваемого уравнения при $\lambda \geqslant 0$, $\rho > 0$, непрерывно зависит от этих аргументов в совокупности.

Доказательство леммы 2. Вследствие аддитивной формы функции (7) достаточно исследовать на минимум разрывную при $\rho=0$ функцию одной переменной

$$\mathcal{F}(\rho)\equiv\lambda\theta(\rho^2)+(\rho-\rho_k^h)^2, \quad \rho\geqslant 0, \quad k=1, \dots, M.$$

Предположим сначала, что $\rho_k^h>0$ для рассматриваемого фиксированного k. Тогда, вводя переменную $x=\rho/\rho_k^h$, получаем

$$\mathcal{F}(\rho)=\{(\rho_k^h)^2\colon \rho=0;\ \varphi(x)\colon \rho>0\}.$$

Здесь

$$\varphi(x)\equiv\lambda(\rho_k^h)^{-2}x^{-2}+(\rho_k^h)^2(x-1)^2, \quad x>0.$$

Легко исследовать, что функция $\varphi(x)$ имеет единственную точку глобального минимума $x_k(\lambda)$, которая находится из уравнения (10) и, следовательно, удовлетворяет неравенству $x_k(\lambda)\geqslant 1$. Кроме того, с учетом (10) получим

$$\begin{aligned}\varphi_{\min}\equiv\inf\{\varphi(x)\colon x>0\}=\varphi[x_k(\lambda)]=\\ =\lambda(\rho_k^h)^{-2}x_k^{-2}(\lambda)+(\rho_k^h)^2[x_k(\lambda)-1]^2=\\ =(\rho_k^h)^2\{\lambda(\rho_k^h)^{-4}x_k^{-2}(\lambda)+[x_k(\lambda)-1]^2\}=\\ =(\rho_k^h)^2\{x_k^2(\lambda)-x_k(\lambda)+[x_k(\lambda)-1]^2\}=(\rho_k^h)^2[2x_k^2(\lambda)-3x_k(\lambda)+1].\end{aligned}$$

Поэтому

$$\begin{aligned}\mathcal{F}_{\min}\equiv\inf\{\mathcal{F}(\rho)\colon \rho\geqslant 0\}=\min\{\mathcal{F}(0), \varphi[x_k(\lambda)]\}=\\ =\min\{\mathcal{F}(0), \mathcal{F}[\rho_k^h x_k(\lambda)]\}=(\rho_k^h)^2\min\{1, 2x_k^2(\lambda)-3x_k(\lambda)+1\}. \quad (11)\end{aligned}$$

Отсюда понятно, что $\mathcal{F}_{\min}=\mathcal{F}[\rho_k^h x_k(\lambda)]$, если $2x_k^2-3x_k\leqslant 0$, т. е. при $1\leqslant x_k(\lambda)\leqslant 3/2$. Это неравенство можно решить относительно λ, учитывая равенство

$$\lambda=(\rho_k^h)^4[x_k^4(\lambda)-x_k^3(\lambda)],$$

вытекающее из (10). Тогда оказывается, что $0\leqslant\lambda\leqslant 27(\rho_k^h)^4/16\equiv\lambda_k$. Тем самым показано, что $\mathcal{F}_{\min}=\mathcal{F}[\rho_k^h x_k(\lambda)]$ при $0\leqslant\lambda\leqslant\lambda_k$. Если же $\lambda\geqslant\lambda_k$, т. е. $x_k(\lambda)\geqslant 3/2$, то вследствие (11) имеем $\mathcal{F}_{\min}=\mathcal{F}(0)$. Итак, формулы (9), (9′) для решения задачи (8) обоснованы в случае, если $\rho_k^h>0$.

Пусть теперь $\rho_k^h=0$. Тогда из вида функции $\mathcal{F}(\rho)$ ясно, что $\mathcal{F}_{\min}=\mathcal{F}(0)$ для каждого $\lambda\geqslant 0$. Поскольку в рассматриваемом случае $\lambda_k=0$, то $\rho_k(\lambda)=0$ при $\lambda>\lambda_k=0$ и формула (9) снова верна. Очевидна также и формула (9′), имеющая здесь вид $\rho_k(0)=0=\rho_k^h$.

Остается установить свойство монотонности, указанное в формулировке леммы. Если $\lambda=0$, то, согласно (9′), имеем $\rho_k(\lambda)=\rho_k^h$, и свойство монотонности выполнено в силу упорядоченности сингулярных чисел ρ_k^h матрицы A_h. Предположим, что $\lambda>0$ и для рассматриваемого $k\geqslant 2$ выполнено неравенство $\rho_k(\lambda)>0$. Тогда из (9) ясно, что $\rho_k^h>0$, так что, согласно свойствам функции $x_k(\lambda)$, выполнено неравенство $x_k(\lambda)>1$. Из (9), (10) вытекает равенство

$$\rho_{k-1}^4(\lambda)/\rho_k^4(\lambda)=[x_{k-1}^4(\lambda)(\rho_{k-1}^h)^4]/[x_k^4(\lambda)(\rho_k^h)^4]=(1-x_k^{-1})/(1-x_{k-1}^{-1})\equiv s.$$

Рассматривая уравнение (10), можно понять, что неравенство $\rho^h_{k-1} \geqslant \rho^h_k > 0$ влечет выполнение неравенства $x_k(\lambda) \geqslant x_{k-1}(\lambda) > 1$. Но тогда $s \geqslant 1$, т. е. $\rho_{k-1}(\lambda) \geqslant \rho_k(\lambda)$. Остается отметить, что это неравенство выполнено и при $\rho_k(\lambda)=0$. Лемма доказана.

Зададим теперь функцию

$$\beta(\lambda) \equiv \sum_{k=1}^{M} [\rho_k(\lambda) - \rho^h_k]^2. \tag{12}$$

Свойства этой функции определяет

Лемма 3. *Функция* (12) *монотонно возрастает при* $0 \leqslant \lambda \leqslant \lambda_1$ *и непрерывна всюду, кроме точек разрыва первого рода* $\lambda = \lambda_p$ $(p=1, \ldots, M)$, *являющихся одновременно точками неоднозначности этой функции. В точках* $\lambda=\lambda_p$ *функция* (12) *принимает два значения*:

$$\beta(\lambda_p+0) = \sum_{k=1}^{p-q(p)} (\rho^h_k)^2 [x_k(\lambda_p)-1]^2 + \sum_{k=p-q(p)+1}^{M} (\rho^h_k)^2,$$
$$\beta(\lambda_p-0) = \beta(\lambda_p+0) - 3q(p)(\rho^h_k)^2/4. \tag{13}$$

Здесь $q(p)$ — *кратность сингулярного числа* ρ^h_p *матрицы* A. *Кроме того,* $\beta(+0)=0$, $\beta(+\infty)=\|A_h\|^2$, $\beta(\lambda)=\beta(+\infty)$ *при* $\lambda \geqslant \lambda_1$.

Доказательство. Функцию (12) можно представить в виде

$$\beta(\lambda) = \sum_{k=1}^{M} \xi_k(\lambda),$$

где функции $\xi_k(\lambda)$ вводятся с учетом (9), (9′) следующим образом:

$$\xi_k(\lambda) \equiv [\rho_k(\lambda) - \rho^h_k]^2 = \{(\rho^h_k)^2[x_k(\lambda)-1]^2\colon 0 \leqslant \lambda \leqslant \lambda_k;\ (\rho^h_k)^2\colon \lambda \geqslant \lambda_k\}.$$

Имеет смысл изучать лишь не равные тождественно нулю функции $\xi_k(\lambda)$, отвечающие тем значениям k $(1 \leqslant k \leqslant M)$, для которых $\rho^h_k > 0$.

Из определения ясно, что функция $\xi_k(\lambda)$ имеет точку неоднозначности $\lambda = \lambda_k$, являющуюся одновременно точкой разрыва первого рода. При этом

$$\xi_k(\lambda_k-0) = (\rho^h_k)^2/4, \quad \xi_k(\lambda_k+0) = (\rho^h_k)^2.$$

Исходя из свойств функции $x_k(\lambda)$, указанных выше, можно вывести непрерывность и монотонное возрастание $\xi_k(\lambda)$ при $0 \leqslant \lambda \leqslant \lambda_k$, а также равенство $\xi_k(0)=0$. График функции $\xi_k(\lambda)$ представлен на рис. 21. Суммируя эти функции при вычислении величины $\beta(\lambda)$, легко установить свойства непрерывности и монотонности функции (12), сформулированные в лемме, а также равенства (13) и соотношения

$$\beta(+0) = \sum_{k=1}^{M} \xi_k(0) = 0, \quad \beta(+\infty) = \sum_{k=1}^{M} (\rho^h_k)^2 = \|A_h\|^2.$$

Из доказанной леммы следует, что при выполнении условия $\|A_h\| > h > 0$ уравнения с монотонной функцией

$$\beta(\lambda) = h^2 \tag{14}$$

имеет единственное решение $\lambda(h) > 0$ (рис. 22).

Теорема 2. *Предположим, что* $\|A_h\| > h$ *и* $\lambda(h)$ — *обычное решение уравнения* (14). *Тогда числа* $\hat\rho_k \equiv \rho_k[\lambda(h)]$ $(k=1, \ldots, M)$, *определяемые согласно* (9), (9′), *дают решение задачи* (1).

Доказательство. В лемме 2 установлено свойство монотонности чисел $\hat\rho_k$: $\hat\rho_1 = \rho_1[\lambda(h)] \geqslant \ldots \geqslant \hat\rho_M = \rho_M[\lambda(h)] \geqslant 0$. Кроме того, из (14) следует равенство

$$\beta[\lambda(h)] = \sum_{k=1}^{M}[\rho_k[\lambda(h)] - \rho_k^h]^2 = \sum_{k=1}^{M}(\hat\rho_k - \rho_k^h)^2 = h^2. \tag{15}$$

Таким образом, числа $\hat\rho_1, \ldots, \hat\rho_M$ удовлетворяют ограничениям условной экстремальной задачи (1). Далее, вектор $(\hat\rho_1, \ldots, \hat\rho_M, \lambda(h))$ представляет собой по построению седловую точку функции Лагранжа (7) этой задачи. Поэтому числа $\hat\rho_1, \ldots, \hat\rho_M$ представляют собой решение задачи (1) (см. [31]).

Теорема 1 определяет способ решения задачи (1), который пригоден, если $\lambda(h)$ представляет собой обычное решение уравнения

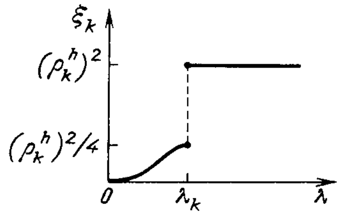

Рис. 21. Вид функции $\xi_k(\lambda)$

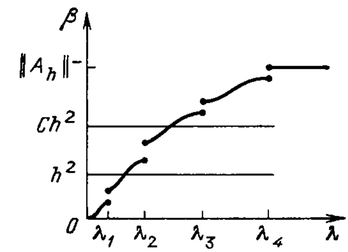

Рис. 22. Примерное поведение невязки $\beta(\lambda)$ для метода м.п.м.

(14). Отметим, что выяснить, имеет ли это уравнение обычное решение или нет, можно, сопоставляя предельные значения $\beta(\lambda_p \pm 0)$ функции (12) в ее точках разрыва λ_p $(p=1, \ldots, M)$ с числом h^2. Уравнение (14) имеет обычное решение, если ни для какого p не выполнено неравенство

$$\beta(\lambda_p - 0) < h^2 < \beta(\lambda_p + 0).$$

Указанные предельные значения вычислены в формуле (13).

Уравнение (14) может не иметь обычного решения. Это реализуется, если число $\lambda(h)$ совпадает с одним из чисел λ_p и, кроме того, выполнены неравенства

$$\beta[\lambda(h) - 0] < h^2, \quad \beta[\lambda(h) + 0] > h^2.$$

Тогда числа $\hat\rho_k$ $(k=1, \ldots, M)$ уже не будут решением задачи (1), так как не выполнено требование (15). В этом случае можно применить второй подход к решению задачи (1). Прежде чем обратиться к нему, рассмотрим проблему устойчивого вычисления ранга матрицы $\bar{A}$ по заданной матрице A_h.

Предположим, что числа $\lambda_1 \geqslant \ldots \geqslant \lambda_M \geqslant 0$ — точки разрыва функции (12), указанные в лемме 2. Формально введем также $\lambda_{M+1}=0$, $\lambda_0=+\infty$.

Теорема 3. *Пусть выполнено условие* $\|A_h\|>h$ *и* $\lambda(h)$ — *решение уравнения* (14). *Тогда*:

а) *при каждом* $h>0$ *существует такой номер* $t=t(h)$ $(0\leqslant t(h)\leqslant M)$, *для которого выполнено неравенство*

$$\lambda_{t(h)+1}\leqslant\lambda(h)<\lambda_{t(h)}; \tag{16}$$

б) *при* $0<h<\|\bar{A}^+\|_*^{-1}/3\equiv h_1(\bar{A})$ *справедливо равенство* $t(h)=\operatorname{Rg}\bar{A}$.

Доказательство. Прежде всего установим существование номера $t(h)$, для которого выполнено (16). Предположим, что такого положительного номера нет. Это означает в силу положительности числа $\lambda(h)$, что $\lambda(h)\geqslant\lambda_k$ для любого $k=1, \ldots, M$, т. е. что $\lambda(h)\geqslant\lambda_1$. Если при этом $\lambda(h)>\lambda_1$, то из условия теоремы и по лемме 3 получим

$$\beta[\lambda(h)]=\beta[\lambda(h)-0]=\beta(+\infty)=\|A_h\|^2>h^2.$$

Но это противоречило бы требованию $\beta[\lambda(h)-0]\leqslant h^2$, которое вытекает из (14). Значит, или $\lambda(h)=\lambda_1$, т. е. $t(h)=0$, или выполнено неравенство (16) с некоторым номером $t(h)$ $(1\leqslant t(h)\leqslant M)$.

Перейдем теперь к доказательству равенства $t(h)=\operatorname{Rg}\bar{A}$ для h, указанных в условиях. При этом необходимо рассмотреть два варианта.

В первом из них для рассматриваемого h выполнено строгое неравенство $\lambda_{t+1}<\lambda(h)<\lambda_t$. Тогда $\lambda(h)$ — точка непрерывности функции (12), и по теореме 2 числа $\hat{\rho}_k=\rho_k[\lambda(h)]$ $(k=1, \ldots, M)$ представляют собой сингулярные числа матрицы $\hat{A}_h$ метода м.п.м. Из формулы (9), взятой для $\lambda=\lambda(h)$, следует, что $\hat{\rho}_k=\rho_k[\lambda(h)]=0$ при $t(h)+1\leqslant k\leqslant M$ и $\hat{\rho}_k>0$ при $1\leqslant k\leqslant t(h)$. Это означает, что $\operatorname{Rg}\hat{A}_h=t(h)$. С другой стороны, по теореме 2.4 выполнение требования $0<h<\|\bar{A}^+\|_*^{-1}/3$ обеспечивает доказываемое равенство $t(h)=\operatorname{Rg}\hat{A}_h=\operatorname{Rg}\bar{A}$.

Во втором варианте для рассматриваемого h выполнено равенство $\lambda(h)=\lambda_{t+1}$. В этом случае число $\lambda(h)$ есть точка разрыва функции (12). Поэтому

$$\beta[\lambda(h)-0]\leqslant h^2\leqslant\beta[\lambda(h)+0]. \tag{17}$$

Кроме того, в силу (13) выполнено равенство

$$\beta[\lambda(h)-0]=\beta[\lambda(h)+0]-\frac{3}{4}q_{t+1}(\rho_{t+1}^h)^2=$$

$$=\sum_{k=1}^{t}(\rho_k^h)^2[x_k[\lambda(h)]-1]^2+\frac{1}{4}q_{t+1}(\rho_{t+1}^h)^2+\sum_{k=t+2}^{M}(\rho_k^h)^2. \tag{18}$$

В нем число q_{t+1} представляет собой кратность сингулярного числа $\rho_{t(h)+1}^h$ матрицы A_h. Из соотношений (17), (18) следует, что

$$\frac{1}{4}q_{t+1}(\rho_{t+1}^h)^2\leqslant\beta[\lambda(h)-0]\leqslant h^2,$$

а это неравенство в свою очередь дает с учетом (18) оценку

$$\beta[\lambda(h)+0]=\sum_{k=1}^{M}[\rho_k[\lambda(h)+0]-\rho_k^h]^2=$$

$$=\beta[\lambda(h)-0]+\frac{3}{4}q_{t+1}(\rho_{t+1}^h)^2\leqslant h^2+3h^2=4h^2. \tag{19}$$

Здесь числа $\rho_k[\lambda(h)+0]\equiv\rho_k^*$ определяются формулой (9).

Далее, из неравенства Минковского и из (19) следует неравенство

$$\left\{\sum_{k=1}^{M}(\rho_k^*-\bar\rho_k)^2\right\}^{1/2}\leqslant\left\{\sum_{k=1}^{M}(\rho_k^*-\rho_k^h)^2\right\}^{1/2}+\left\{\sum_{k=1}^{M}(\rho_k^h-\bar\rho_k)^2\right\}^{1/2}\leqslant$$

$$\leqslant\beta^{1/2}[\lambda(h)+0]+h\leqslant 3h. \tag{20}$$

Числа ρ_k^* $(k=1, ..., M)$ дают по лемме 2 решение задачи (8) для $\lambda=\lambda(h)$. Поэтому с учетом неравенства (17) можно сделать следующую выкладку:

$$\lambda(h)\sum_{k=1}^{M}\theta[(\rho_k^*)^2]+h^2\leqslant\lambda(h)\sum_{k=1}^{M}\theta[(\rho_k^*)^2]+\beta[\lambda(h)+0]=$$

$$=\lambda(h)\sum_{k=1}^{M}\theta[(\rho_k^*)^2]+\sum_{k=1}^{M}(\rho_k^*-\rho_k^h)^2=M^{\lambda(h)}[\rho_1^*, ..., \rho_M^*]\leqslant$$

$$\leqslant M^{\lambda(h)}[\bar\rho_1, ..., \bar\rho_M]=\lambda(h)\sum_{k=1}^{M}\theta(\bar\rho_k^2)+\sum_{k=1}^{M}(\bar\rho_k-\rho_k^h)^2\leqslant$$

$$\leqslant\lambda(h)\|\bar A^+\|_*^2+\|\bar A-A_h\|^2\leqslant\lambda(h)\|\bar A^+\|_*^2+h^2.$$

Здесь использованы также соотношения (2.8), (2.9). Отсюда

$$\sum_{k=1}^{M}\theta[(\rho_k^*)^2]\leqslant\|\bar A^+\|_*^2. \tag{21}$$

Введем теперь матрицу $A_*^h\equiv\bar U D_*^h \bar V^T$, определяемую с помощью входящих в сингулярное разложение матрицы $\bar A$ ортогональных матриц $\bar U$, $\bar V$, а также с помощью диагональной матрицы $D_*^h\equiv\operatorname{diag}(\rho_1^*, ..., \rho_M^*)\in\mathfrak{A}$, которая содержит сингулярные числа матрицы A_*^h. Из формулы (9), взятой при $\lambda=\lambda(h)-0=\lambda_t-0$, следует, что $\rho_k^*=\rho_k^h x_k(\lambda_{t+1})>0$ для $k=1, ..., t(h)$ и $\rho_k^*=0$ для $k=t(h)+1, ..., M$. Поэтому $\operatorname{Rg}A_*^h=t(h)$.

Неравенства (20), (21) можно переписать в следующей форме:

$$\|A_*^h-\bar A\|^2=\|\bar U(D_*^h-\bar D)\bar V^T\|^2=\|D_*^h-\bar D\|^2=\sum_{k=1}^{M}(\rho_k^*-\bar\rho_k)^2\leqslant(3h)^2,$$

$$\|(A_*^h)^+\|_*^2=\sum_{k=1}^{M}\theta[(\rho_k^*)^2]\leqslant\|\bar A^+\|_*^2.$$

Эти неравенства являются условиями следствия 2.5 и приводят к равенству $t(h)=\operatorname{Rg}A_*^h=\operatorname{Rg}\bar A$, справедливому при $0<h<\|\bar A^+\|_*^{-1}/3$. Теорема доказана.

Из доказательства теоремы 3 следует, что равенство $\lambda(h)=\lambda_1$ невозможно при $0<h<h_1(\bar{A})$, так как иначе выполнялось бы равенство $\mathrm{Rg}\,\bar{A}=t(h)=0$, т. е. $\bar{A}=0$. Последнее запрещено условием $\|A_h\|>h\geqslant\|\bar{A}-A_h\|$. Таким образом, при достаточно малых h условие (16) реализуется при положительном номере $t(h)$.

Теорема 3 определяет устойчивый при $h\to 0$ метод нахождения ранга матрицы $\bar{A}$ по заданной приближенной матрице A_h. Этот метод основан на вычислении решения $\lambda(h)>0$ уравнения (14) с последующим нахождением номера $t(h)$ из условия (16) путем сравнения чисел $\lambda_k=27(\rho_k^h)^4/16$ $(k=1, \ldots, M)$, $\lambda_{M+1}=0$, $\lambda_0=+\infty$, $\lambda(h)$.

Как уже было отмечено, теоремы 1, 2 позволяют найти матрицу метода м.п.м. в случае, если уравнение (14) имеет обычное решение $\lambda(h)>0$. Если же это требование не выполнено, можно найти некоторый аналог матрицы метода м.п.м., который также позволяет строить устойчивые приближения к нормальному псевдорешению $\bar{z}$ СЛАУ (1.1). Делается это на основе анализа доказательства теоремы 3 в той его части, где рассматривается вариант $\lambda(h)=\lambda_{t(h)+1}$.

Итак, пусть уравнение (14) имеет обобщенное решение $\lambda(h)=\lambda_{t(h)+1}$, где $t(h)=\mathrm{Rg}\,\bar{A}$ при $0<h<h_1(\bar{A})$. Используем введенные при доказательстве теоремы 3 числа ρ_k^*:

$$\rho_k^*=\{\rho_k^h x_k[\lambda_{t(h)+1}]\colon k=1, \ldots, t(h);\quad 0\colon k=t(h)+1, \ldots, M\}$$

и образуем матрицу $A(h)=U_h D(h) V_h^T$, где $D(h)\equiv\mathrm{diag}(\rho_1^*, \ldots, \rho_M^*)\in\mathfrak{A}$. Тогда $A^+(h)=V_h D^+(h) U_h^T$, причем $D^+(h)=\mathrm{diag}[\theta(\rho_1^*), \ldots, \theta(\rho_M^*)]\in\mathfrak{A}^*$. Образуем также элемент $\hat{z}_\eta=A^+(h)u_\sigma$.

Теорема 4. *При $\eta=(h, \sigma)\to 0$ приближенные решения $\hat{z}_\eta$ сходятся к $\bar{z}$. При $0<h<h_1(\bar{A})$ справедлива оценка скорости сходимости*

$$\|\hat{z}_\eta-\bar{z}\|\leqslant\|\bar{A}^+\|_*\cdot[\sigma+3h\|\bar{A}^+\|_*\cdot\|\bar{u}\|/(1-3h\|\bar{A}^+\|_*)^3].$$

Доказательство. На основании оценки (19) получим

$$\|A(h)-A_h\|^2=\|U_h[D(h)-D_h]V_h^T\|^2=$$

$$=\|D(h)-D_h\|^2=\sum_{k=1}^{M}(\rho_k^*-\rho_k^h)^2=\beta[\lambda(h)+0]\leqslant 4h^2.$$

Поэтому

$$\|A(h)-\bar{A}\|\leqslant\|A(h)-A_h\|+\|A_h-\bar{A}\|\leqslant 3h. \tag{22}$$

Далее в силу неравенства (21) и на основании равенства (2.9) справедливо соотношение

$$\|A^+(h)\|_*^2=\sum_{k=1}^{M}\theta[(\rho_k^*)^2]\leqslant\|\bar{A}^+\|_*^2. \tag{23}$$

Из неравенств (22), (23), рассуждая, как при доказательстве теоремы 2.6, получаем при $0<h<h_1(\bar{A})$ оценку

$$\|A^+(h)-\bar{A}^+\|_*\leqslant 3h\|\bar{A}^+\|_*^2/(1-3h\|\bar{A}^+\|_*)^3,$$

из которой по аналогии с соотношением (2.16) вытекает искомая оценка скорости сходимости:

$$\|z_\eta-\bar z\|\leqslant\|\bar A^+\|_*\sigma+\|A^+(h)-\bar A^+\|_*\cdot\|\bar u\|\leqslant$$
$$\leqslant\|\bar A^+\|_*\cdot\left[\sigma+3h\|\bar A^+\|_*\cdot\|\bar u\|/\left(1-3h\|\bar A^+\|_*\right)^3\right].$$

Теорема доказана.

Итак, матрица $A(h)$ может заменить матрицу метода м.п.м. для построения приближений к $\bar z$. Приближенное решение $\hat z_\eta$, получаемое с помощью этой матрицы, имеет оптимальный порядок точности $O(h+\sigma)$ вне зависимости от того, совместна система (1.1) или нет. Однако такая $A(h)$ не удовлетворяет неравенству $\|A-A_h\|\leqslant h$.

Обратимся теперь ко второму способу нахождения решения задачи (1). Он основан на использовании экстремальной задачи: при фиксированном r $(1\leqslant r\leqslant M)$ найти такие числа $\check\rho_1, \ldots, \check\rho_r>0$, для которых

$$\sum_{k=1}^r\check\rho_k^{-2}=\inf\left\{\sum_{k=1}^r\rho_k^{-2}:\ \rho_1, \ldots, \rho_r>0,\right.$$
$$\left.\sum_{k=1}^r(\rho_k-\rho_k^h)^2+\sum_{k=r+1}^M(\rho_k^h)^2=h^2\right\}. \qquad (24)$$

Для решения задачи (24) введем функцию Лагранжа

$$\mathscr{L}^\mu(\rho_1, \ldots, \rho_r)=\mu\sum_{k=1}^r\rho_k^{-2}+\sum_{k=1}^r(\rho_k-\rho_k^h)^2, \qquad (25)$$
$$\mu\geqslant 0,\quad \rho_1, \ldots, \rho_r>0.$$

Лемма 4. *Пусть $r=\mathrm{Rg}\,\bar A$, $0<h<h_0(\bar A)$. Тогда задача* (24) *имеет единственное решение $\check\rho_1, \ldots, \check\rho_r>0$. Оно обладает свойством монотонности: $\check\rho_1\geqslant\ldots\geqslant\check\rho_r>0$. Числа $\check\rho_k$ $(k=1, \ldots, r)$ определяется формулой $\check\rho_k=\rho_k^h x_k[\mu(h)]$, где $x_k(\mu)$ — положительное решение уравнения*

$$x^4-x^3=\mu(\rho_k^h)^{-4},\quad \mu\geqslant 0, \qquad (26)$$

а число $\mu(h)\geqslant 0$ — единственный корень уравнения

$$\varepsilon(\mu)\equiv\sum_{k=1}^r(\rho_k^h)^2[x_k(\mu)-1]^2+\sum_{k=r+1}^M(\rho_k^h)^2=h^2 \qquad (27)$$

с непрерывной и монотонно возрастающей при $\mu\geqslant 0$ функцией $\varepsilon(\mu)$.

Доказательство. Необходимые условия экстремума функции Лагранжа (25) дают $\rho_k^3(\rho_k-\rho_k^h)=\mu$ $(k=1, \ldots, r)$. Это уравнение после замены $x_k=\rho_k/\rho_k^h$ приводится к виду (26). При этом используется следствие 2.4: $\mathrm{Rg}\,A_h\geqslant\mathrm{Rg}\,\bar A=r$, из которого следует неравенство $\rho_k^h>0$ $(k=1, \ldots, r)$. Выше уже отмечалось, что уравнение (26) имеет единственное положительное решение $x_k(\mu)\geqslant 1$ при $\mu\geqslant 0$. Поэтому функция (25) обладает единственной стационарной точкой $\rho_k(\mu)=\rho_k^h x_k(\mu)$ $(k=1, \ldots, r)$, являющейся точкой ее минимума, так как

$$d^2\mathscr{L}^\mu[\rho_1(\mu), \ldots, \rho_r(\mu)]=2\sum_{k=1}^r[1+3\mu\rho_k^{-4}(\mu)](d\rho_k)^2>0.$$

Найдем множитель Лагранжа $\mu\geqslant 0$ из условия связи задачи (24), имеющего при $\rho_k=\rho_k(\mu)$ вид уравнения (27). При этом заметим, что функция $\varepsilon(\mu)$ монотонно возрастает и непрерывна при $\mu\geqslant 0$ вследствие аналогичных свойств функций $x_k(\mu)$ $(k=1, ..., r)$. Из (26) ясно, что $x_k(0)=1$, $x_k(+\infty)=+\infty$, так что $\varepsilon(+\infty)=+\infty$ и

$$\varepsilon(0)=\sum_{k=r+1}^{M}(\rho_k^h)^2\leqslant\sum_{k=r+1}^{M}(\rho_k^h)^2+\sum_{k=1}^{r}(\bar{\rho}_k-\rho_k^h)^2\leqslant\|\bar{A}-A_h\|^2\leqslant h^2.$$

Установленные свойства функции $\varepsilon(\mu)$ гарантируют однозначную разрешимость уравнения (27). Но тогда числа $\check{\rho}_1, ..., \check{\rho}_r$, $\mu(h)$ определяют седловую точку функции Лагранжа (25), и поэтому значения $\check{\rho}_1, ..., \check{\rho}_r$ представляют собой решение задачи (24).

Покажем, что это решение единственно. Пусть это не так и существует также другое решение $\tilde{\rho}_1, ..., \tilde{\rho}_r>0$ задачи (24). Тогда

$$\sum_{k=1}^{r}\tilde{\rho}_k^{-2}=\sum_{k=1}^{r}\check{\rho}_k^{-2},$$

$$\sum_{k=1}^{r}(\tilde{\rho}_k-\rho_k^h)^2=\sum_{k=1}^{r}(\check{\rho}_k-\rho_k^h)^2=h^2-\sum_{k=r+1}^{M}(\rho_k^h)^2,$$

откуда следует равенство $\mathscr{L}^{\mu(h)}(\tilde{\rho}_1, ..., \tilde{\rho}_r)=\mathscr{L}^{\mu(h)}(\check{\rho}_1, ..., \check{\rho}_r)$. Но, с другой стороны, в силу единственности точки глобального минимума функции (25) справедливо строгое неравенство

$$\mathscr{L}^{\mu(h)}(\check{\rho}_1, ..., \check{\rho}_r)<\mathscr{L}^{\mu(h)}(\tilde{\rho}_1, ..., \tilde{\rho}_r).$$

Полученное противоречие доказывает единственность решения задачи (24). Свойство монотонности решения доказывается, как в лемме 2.

Следующая теорема устанавливает связь между решениями задач (1) и (24).

Теорема 5. *Если при* $0\leqslant h<h_0(\bar{A})$ *выполнено условие* $\|A_h\|>h$, *то задача* (1) *имеет единственное решение. Оно имеет вид* $(\hat{\rho}_1, ..., \hat{\rho}_M)=$ $=(\check{\rho}_1, ..., \check{\rho}_r, 0, ..., 0)$, *где* $(\check{\rho}_1, ..., \check{\rho}_r)$ — *решение задачи* (24), *найденное в лемме* 4.

Доказательство. При $0\leqslant h<h_0(\bar{A})$ по теореме 2.4 ранг любой матрицы метода м.п.м. равен r. Следовательно, такой же ранг должна иметь и матрица $\hat{A}_h$ из теоремы 1, а значит, и матрица $\hat{D}_h$, содержащая на своей диагонали числа $\hat{\rho}_1, ..., \hat{\rho}_M$ — решения задачи (1). Поэтому $\hat{\rho}_k>0$ при $1\leqslant k\leqslant r$ и $\hat{\rho}_k=0$ при $r<k\leqslant M$. Из (1) ясно, что для таких чисел $\hat{\rho}_k$ выполнены условия

$$\sum_{k=1}^{M}(\hat{\rho}_k-\rho_k^h)^2=\sum_{k=1}^{r}(\hat{\rho}_k-\rho_k^h)^2+\sum_{k=r+1}^{M}(\rho_k^h)^2=h^2,$$

$$\sum_{k=1}^{M}\theta(\hat{\rho}_k^2)=\sum_{k=1}^{r}\hat{\rho}_k^{-2}\leqslant\sum_{k=1}^{r}\check{\rho}_k^{-2}.$$

Последнее неравенство вытекает из того факта, что числа $(\check{\rho}_1, ..., \check{\rho}_r, 0, ..., 0)$ удовлетворяют ограничениям условной экстремальной задачи (1) (см. задачу (24) и лемму 4). Из этих условий поэтому можно сделать вывод, что числа $\hat{\rho}_1, \ldots, \hat{\rho}_r$ представляют собой решение задачи (24). В силу его единственности будут справедливы равенства $\hat{\rho}_k = \check{\rho}_k$ $(k=1, \ldots, r)$. Теорема доказана.

Таким образом, при $0<h<h_0(\bar{A})$, находя число r по теореме 3, можно определить решение задачи (1) по формулам, указанным в лемме 4.

Из теоремы 5 получается следующая важная

Теорема 6. *При выполнении условий* $0 \leqslant h \leqslant h_0(\bar{A})$, $\|A_h\|>h$ *любая матрица метода м.п.м. имеет одни и те же сингулярные числа* $\hat{\rho}_1, \ldots, \hat{\rho}_M$ — *единственное решение задачи* (1).

Доказательство. Возьмем произвольную матрицу метода м.п.м. $\tilde{A}_h$ — решение задачи (2.1). Обозначим ее сингулярные числа как $\tilde{\rho}_1 \geqslant \ldots \geqslant \tilde{\rho}_M \geqslant 0$. В теореме 1 показано, что матрица $\hat{A}_h$ с сингулярными числами $\hat{\rho}_1 \geqslant \ldots \geqslant \hat{\rho}_M \geqslant 0$ также представляет собой решение задачи (2.1). Поэтому из (2.1), (2.8), (2.9) вытекают соотношения

$$\|\tilde{A}_h^+\|_*^2 = \sum_{k=1}^{M} \theta(\tilde{\rho}_k^2) = \|\hat{A}_h^+\|_*^2 = \sum_{k=1}^{M} \theta(\hat{\rho}_k^2), \tag{28}$$

$$\|\tilde{A}_h - A_h\|^2 = h^2 \geqslant \sum_{k=1}^{M} (\tilde{\rho}_k - \rho_k^h)^2. \tag{29}$$

Из (29) ясно, что числа $\tilde{\rho}_k$ $(k=1, \ldots, M)$ удовлетворяют ограничениям условной экстремальной задачи (2), эквивалентной по лемме 1 задаче (1). Из (28) следует также, что эти числа реализуют точную нижнюю грань в задаче (2) и, значит, в задаче (1). Поэтому $\tilde{\rho}_1, \ldots, \tilde{\rho}_M$ — решение задачи (1). В силу его единственности (при выполнении условий теоремы) тогда оказывается, что $\tilde{\rho}_k = \hat{\rho}_k$ $(k=1, \ldots, M)$. Теорема доказана.

В заключение рассмотрим некоторые примеры.

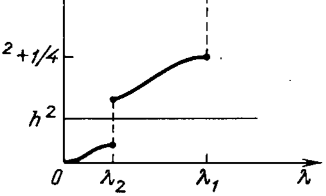

Рис. 23. Иллюстрация к примеру 2

Пример 1. *Нахождение сингулярных чисел матрицы метода м.п.м.* Пусть сингулярные числа точной матрицы $\bar{A}$ имеют вид $\rho_1=1$, $\rho_k=0$ $(k=2, \ldots, M)$, и пусть нам известны сингулярные числа приближенной матрицы A_h: $\rho_1^h=1$, $\rho_2^h=h_2, \ldots, \rho_M^h=h_M$, причем $h_k \geqslant 0$, $h_2^2+\ldots+h_M^2 \leqslant h^2$. Используя результат теоремы 5, можно тогда утверждать, что при $0<h<1/2$ сингулярные числа $\hat{\rho}_k$ матрицы метода м.п.м. для $k=2, \ldots, M$ равны нулю. Поскольку числа $\hat{\rho}_k$ $(k=1, \ldots, M)$ удовлетворяют ограничениям задачи (1), то

$$\sum_{k=1}^{M} (\hat{\rho}_k - \rho_k^h)^2 = (\hat{\rho}_1 - 1)^2 + \sum_{k=2}^{M} h_k^2 = h^2.$$

Тогда задача (1) сводится к выбору из двух чисел, подчиненных этому равенству, такого, для которого меньше число $\theta(\hat\rho_1^2)=\hat\rho_1^{-2}$. Поэтому

$$\hat\rho_1=1+\left[h^2-\sum_{k=2}^{M}h_k^2\right]^{1/2}.$$

Пример 2. *Устойчивое определение ранга матрицы $\bar A$.* Пусть

$$\bar A=\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad A_h=\begin{bmatrix}1&0\\0&h\end{bmatrix},\quad 0<h<1/2.$$

В этом случае $\rho_1^h=1$, $\rho_2^h=h$, $\lambda_1=27/16$, $\lambda_2=27h^4/16$. Вычислим функцию $\beta(\lambda)$ согласно лемме 3:

$$\beta(\lambda)=[\rho_1(\lambda)-\rho_1^h]^2+[\rho_2(\lambda)-\rho_2^h]^2=\{[x_1(\lambda)-1]^2+$$
$$+h^2[x_2(\lambda)-1]^2\colon 0\leqslant\lambda\leqslant\lambda_2;\ [x_1(\lambda)-1]^2+h^2\colon \lambda_2\leqslant\lambda\leqslant\lambda_1;\ 1+h^2\colon \lambda\geqslant\lambda_1\}.$$

График этой функции приведен на рис. 23. В данном примере уравнение (14) имеет обобщенное решение $\lambda(h)=\lambda_2$, так как из (13) следуют неравенства

$$\beta(\lambda_2-0)=\frac{h^2}{4}+[x_1(\lambda_2)-1]^2\leqslant\frac{h^2}{4}+\lambda_2^2=\frac{h^2}{4}+\left(\frac{27h^4}{16}\right)^2<h^2,$$

$$\beta(\lambda_2+0)=\beta(\lambda_2-0)+3h^2/4>h^2.$$

Здесь использована оценка $0\leqslant 1-x_1(\lambda)\leqslant\lambda$, вытекающая из уравнения $x_1^4-x_1^3=\lambda$ при условии, что $1\leqslant x_1\leqslant 3/2$. Таким образом, выполнено неравенство $\lambda_2\leqslant\lambda(h)<\lambda_1$, т. е. по теореме 3 $\operatorname{Rg}\bar A=t(h)=1$ при $0<h<1/2$.

§ 4. Принцип м.п.м.

Рассмотрим несколько иной подход к решению задач (2.1), (2.4) метода м.п.м. Для этого используем метод неопределенных множителей Лагранжа. С этой целью введем функционал

$$M^\lambda[A]=\lambda\|A^+\|_*^2+\|A-A_h\|^2,\quad \lambda>0,\quad A\in\mathfrak{A}_0, \tag{1}$$

и для него рассмотрим экстремальную задачу: при фиксированном $\lambda>0$ найти такую матрицу-экстремаль $A_\lambda\equiv A_{h\lambda}\in\mathfrak{A}_0$, что

$$M^\lambda[A_\lambda]=\inf\{M^\lambda[A]\colon A\in\mathfrak{A}_0\}\equiv\varphi(\lambda). \tag{2}$$

Теорема 1. *При каждом $\lambda>0$ задача* (2) *разрешима.*

Доказательство. Пусть $\{A_n\}$ — какая-либо минимизирующая последовательность для задачи (2) при фиксированном $\lambda>0$: $A_n\in\mathfrak{A}_0$, $M^\lambda[A_n]\to\varphi(\lambda)$ при $n\to\infty$. Тогда найдется такое число $\varepsilon(n_0)$, что при $n\geqslant n_0$ выполнены оценки

$$\|A_n\|\leqslant\|A_h\|+\|A_n-A_h\|\leqslant\|A_h\|+\{M^\lambda[A_n]\}^{1/2}\leqslant$$
$$\leqslant\|A_h\|+\{\varphi(\lambda)+\varepsilon(n_0)\}^{1/2}\equiv M_1=\mathrm{const},$$
$$\|A_n^+\|_*\leqslant\{M^\lambda[A_n]/\lambda\}^{1/2}\leqslant\{[\varphi(\lambda)+\varepsilon(n_0)]/\lambda\}^{1/2}\equiv M_2=\mathrm{const}.$$

Эти неравенства по следствию 2.1 и лемме 2.1 обеспечивают существование подпоследовательности $\{A_{n_l}\}_{l=1}^{\infty}$ такой, что $A_{n_l}\to A_0\in\mathfrak{A}_0$, $A_{n_l}^+\to A_0^+$ при $l\to\infty$. Поэтому

$$\lim_{l\to\infty} M^\lambda[A_{n_l}]=\lim_{l\to\infty}\{\lambda\|A_{n_l}^+\|_*^2+\|A_{n_l}-A_h\|^2\}=$$

$$=\lambda\|A_0^+\|_*^2+\|A_0-A_h\|^2=\varphi(\lambda)=\inf\{M^\lambda[A]: A\in\mathfrak{A}_0\},$$

т. е. A_0 — решение задачи (2) при рассматриваемом λ.

Ниже будет показано, что решение задачи (2) может и не быть единственным (по крайней мере для некоторых чисел λ).

Введем при фиксированном h многозначные функции

$$\beta(\lambda)\equiv\|A_\lambda-A_h\|^2,\quad \gamma(\lambda)\equiv\|A_\lambda^+\|_*^2. \tag{3}$$

Эти функции аналогичны введенным в § 1.6, 2.2 вспомогательным функциям $\beta(\alpha)$, $\gamma(\alpha)$. Используя методы из § 1.6, можно установить основные свойства функций (3). В частности, для $\beta(\lambda)$ справедлива

Лемма 1. *Функция $\beta(\lambda)$ монотонно не убывает и однозначна при $\lambda>0$ всюду, за исключением, быть может, не более чем счетного множества точек разрыва первого рода, являющихся точками многозначности этой функции. Если λ_0 — точка разрыва функции $\beta(\lambda)$, то множество экстремалей функционала* (1) *для этого значения λ содержит по крайней мере две матрицы $A_{\lambda_0\pm 0}$ такие, что $\beta(\lambda_0\pm 0)=\|A_{\lambda_0\pm 0}-A_h\|^2$. Если, кроме того, $0\in\mathfrak{A}_0$, $A_h\in\mathfrak{A}_0$, то $\beta(+0)=0$, $\beta(+\infty)=\|A_h\|^2$.*

Отметим, что при доказательстве существования матриц $A_{\lambda_0\pm 0}$ по аналогии с леммой 1.5 вместо свойств стабилизатора Ω (полунепрерывность снизу в некоторой топологии, компактность в этой топологии множеств уровня стабилизатора) используются результаты леммы 2.1 и следствия 2.1.

Из лемм 1 и 1.6.7 следует

Теорема 2. *Предположим, что $0\in\mathfrak{A}_0$, $A_h\in\mathfrak{A}_0$, $\|A_h\|^2>Ch^2>0$, где $C=\mathrm{const}>1$. Тогда уравнение с монотонной функцией*

$$\beta(\lambda)=Ch^2 \tag{4}$$

имеет решение $\lambda(h)>0$.

Используем значение $\lambda=\lambda(h)$, найденное из (4), и выберем из множества решений задачи (2) при $\lambda=\lambda(h)$ матрицу $A_{\lambda(h)}\equiv A_{\lambda(h)-0}$, для которой вследствие (4) будет выполнено неравенство

$$\beta[\lambda(h)-0]=\|A_{\lambda(h)}-A_h\|^2\leqslant Ch^2. \tag{5}$$

Нетрудно видеть, что найденная таким образом матрица $A_{\lambda(h)}$ построена по схеме алгоритма обобщенного принципа невязки для решения на множестве $\mathfrak{A}_0$ матриц A совместного операторного уравнения $B(A)=\bar{A}$. В этом уравнении B — единичный оператор из $\mathfrak{A}$ в $\mathfrak{A}$, заданный точно, а правая часть $\bar{A}$ известна приближенно, так как вместо нее задана A_h. Вместо функционала Ω берется $\|A^+\|_*^2$

Из результатов § 2.6 тогда получается

Теорема 3. *Справедлива оценка*

$$\| A^+_{\lambda(h)} \|^2_* \leqslant \frac{C}{C-1} \| \bar{A}^+ \|^2_* . \tag{6}$$

Отсюда следует

Теорема 4. *Пусть выполнены условия теоремы 2. Тогда* $A^+_{\lambda(h)} \to \bar{A}^+$ *при* $h \to 0$.

Действительно, из (5) ясно, что $A_{\lambda(h)} \to \bar{A}$ при $h \to 0$. Из этой сходимости и оценки (6) по лемме 2.1 тогда и получим искомую сходимость.

Введем теперь элементы $\hat{z}_\eta = A^+_{\lambda(h)} u_\sigma$. Используя неравенство типа (2.6), а также сходимость, установленную в теореме 4, можно обосновать следующую теорему.

Теорема 5. *Предположим, что для каждого* $h > 0$ *выполнены условия теоремы 2. Тогда* $\hat{z}_\eta \to \bar{z}$ *при* $\eta = (h, \sigma) \to 0$.

Будем говорить, что элемент $\hat{z}_\eta$ *построен по принципу м.п.м.*, а матрица $A_{\lambda(h)}$ будет называться *матрицей принципа м.п.м.*

В случае, если $\mathfrak{A}_0 = \mathfrak{A}$, принцип м.п.м. может быть значительно упрощен, а его матрица вычислена непосредственно. Рассмотрим этот случай.

Теорема 6. *Если* $\mathfrak{A}_0 = \mathfrak{A}$, *то все решения задачи* (2) *при данном* $\lambda > 0$ *имеют вид* $\hat{A}_\lambda = \hat{U} \hat{D}_\lambda \hat{V}^T$. *Здесь диагональная матрица* $\hat{D}_\lambda = \mathrm{diag}[\rho_1(\lambda), \ldots, \rho_M(\lambda)] \in \mathfrak{A}$ *содержит сингулярные числа матрицы* $\hat{A}_\lambda$, *задаваемые формулой* (3.9), *а ортогональные матрицы* $\hat{U}$, $\hat{V}$ ($\hat{U} \in \mathfrak{A}_m$, $\hat{V} \in \mathfrak{A}_n$) *удовлетворяют равенству*

$$\| \hat{U} \hat{D}_\lambda \hat{V}^T - A_h \| = \| \hat{D}_\lambda - D_h \| .$$

В качестве $\hat{U}$, $\hat{V}$ *можно, в частности, взять матрицы* U_h, V_h, *входящие в сингулярное разложение для* A_h.

Доказательство. Прежде всего покажем, что матрица $\hat{A}_\lambda$, определенная в условиях теоремы, является решением задачи (2). По условию на матрицы $\hat{U}$, $\hat{V}$ и в силу равенства (2.9) справедливо соотношение (см. (3.8))

$$M^\lambda[\hat{A}_\lambda] = \lambda \| \hat{A}^+_\lambda \|^2_* + \| \hat{A}_\lambda - A_h \|^2 = \lambda \| \hat{D}^+_\lambda \|^2_* + \| \hat{D}_\lambda - D_h \|^2 =$$
$$= \lambda \sum_{k=1}^{M} \theta[\rho^2_k(\lambda)] + \sum_{k=1}^{M} [\rho_k(\lambda) - \rho^h_k]^2 = \mathscr{Z}^\lambda[\rho_1(\lambda), \ldots, \rho_M(\lambda)]. \tag{7}$$

Поскольку по лемме 3.2 числа $\rho_k(\lambda)$ ($k = 1, \ldots, M$), и только они, реализуют точную нижнюю грань (3.8), то

$$M^\lambda[\hat{A}_\lambda] = \mathscr{Z}^\lambda[\rho_1(\lambda), \ldots, \rho_M(\lambda)] < \mathscr{Z}^\lambda[\rho_1, \ldots, \rho_M]$$

для любого набора чисел $(\rho_1, \ldots, \rho_M)$, где $\rho_1 \geqslant \ldots \geqslant \rho_M \geqslant 0$, отличного от $[\rho_1(\lambda), \ldots, \rho_M(\lambda)]$. Учитывая далее вид (3.7) функции $\mathscr{Z}^\lambda$, а также соотношения (2.8), (2.9), получаем

$$M^\lambda[\hat{A}_\lambda] < \mathscr{Z}^\lambda(\rho_1, \ldots, \rho_M) = \lambda \sum_{k=1}^{M} \theta(\rho^2_k) + \sum_{k=1}^{M} (\rho_k - \rho^h_k)^2 \leqslant$$
$$\leqslant \lambda \| A^+ \|^2_* + \| A - A_h \|^2 = M^\lambda[A]. \tag{8}$$

Это неравенство справедливо для любой матрицы A, имеющей сингулярные числа $\rho_1, \ldots, \rho_M$. В силу произвольности последних неравенство

$$M^\lambda[\hat{A}_\lambda] < M^\lambda[A]$$

выполнено для всякой матрицы $A \in \mathfrak{A}$, имеющей сингулярные числа, отличные от $\rho_1(\lambda), \ldots, \rho_M(\lambda)$. Это и означает, что $\hat{A}_\lambda$ — решение задачи (2) при $\mathfrak{A}_0 = \mathfrak{A}$.

Остается доказать, что задача (2) не будет содержать других решений, кроме матриц $\hat{A}_\lambda$. Действительно, в силу соотношений (7), (8) любое решение задачи (2) с $\mathfrak{A}_0 = \mathfrak{A}$ имеет одни и те же сингулярные числа $\rho_k(\lambda)$ ($k=1, \ldots, M$). Если теперь предположить, что задача (2) имеет решение $A_* = \hat{U}\hat{D}_\lambda \hat{V}^T$, в котором ортогональные матрицы $\hat{U}$, $\hat{V}$ не удовлетворяют указанному в формулировке теоремы условию, то в силу (2.8)

$$\| A_* - A_h \| = \| \hat{U}\hat{D}_\lambda \hat{V}^T - A_h \| > \| \hat{D}_\lambda - D_h \|.$$

Это неравенство вместе с (7) приводит к оценке

$$M^\lambda[\hat{A}_\lambda] = \lambda \| \hat{D}_\lambda^+ \|_*^2 + \| D_\lambda - D_h \|^2 < \lambda \| \hat{V}\hat{D}_\lambda^+ \hat{U}^T \|_*^2 + \| A_* - A_h \|^2 =$$

$$= \lambda \| A_*^+ \|_*^2 + \| A_* - A_h \|^2 = M^\lambda[A_*].$$

Отсюда ясно, что матрица A_* не является решением задачи (2). Теорема доказана.

Из теоремы 6 и формулы (3.9) можно сделать вывод о том, что при каждом $\lambda = \lambda_k = 27(\rho_k^h)^4/16$ ($k=1, \ldots, M$) задача (2) с $\mathfrak{A}_0 = \mathfrak{A}$ имеет по меньшей мере два решения, отличающиеся k-м сингулярным числом: $\rho_k(\lambda_k) = 3\rho_k^h/2$, $\rho_k(\lambda_k) = 0$.

Результат теоремы 6 позволяет более детально исследовать функцию $\beta(\lambda)$. Действительно, по теореме 6 и в силу леммы 3.3 имеет место представление

$$\beta(\lambda) = \sum_{k=1}^{M} [\rho_k(\lambda) - \rho_k^h]^2,$$

так что введенная выше функция $\beta(\lambda)$ совпадает с функцией из (3.12). Поэтому к свойствам функции $\beta(\lambda)$, перечисленным в лемме 1, можно добавить при $\mathfrak{A}_0 = \mathfrak{A}$ дополнительные ее свойства из леммы 3.3. Функция $\beta(\lambda)$ имеет конечное число точек разрыва $\lambda_1 \geqslant \ldots \geqslant \lambda_M \geqslant 0$, где $\lambda_k = 27(\rho_k^h)^4/16$ ($k=1, \ldots, M$); она монотонно возрастает при $0 \leqslant \lambda \leqslant \lambda_1$, и $\beta(\lambda) = \beta(+\infty) = \| A_h \|^2$ при $\lambda \geqslant \lambda_1$.

Из этих свойств вытекает единственность решения $\lambda(h) > 0$ уравнения $\beta(\lambda) = h^2$ при $\| A_h \| > h > 0$. Используем найденное таким образом число $\lambda(h)$, являющееся решением уравнения (3.14), и построим матрицы вида $\hat{A}_{\lambda(h)} = \hat{U}\hat{D}_{\lambda(h)}\hat{V}^T$ по теореме 6. Из теорем 3.1, 3.2, 3.5, 6 непосредственно получается

Теорема 7. *Пусть при* $0 < h < h_0(\bar{A})$ *выполнено условие* $\| A_h \| > h$ *и число* $\lambda(h)$ *есть обычное решение уравнения* (4) *с* $C=1$. *Тогда матрицы* $\hat{A}_{\lambda(h)}$ *являются одновременно матрицами метода м.п.м. и принципа м.п.м.*

Таким образом, устанавливается связь метода и принципа м.п.м. Эта связь аналогична связи о.м.н. и о.п.н. (см. § 2.5).

Отметим, что матрица $\hat{D}_{\lambda(h)}$ содержит сингулярные числа $\hat{\rho}_k = \rho_k[\lambda(h)]$ (см. теорему 3.2). В качестве матриц $\hat{U}$, $\hat{V}$ можно принять U_h, V_h.

§ 5. Оптимальные свойства метода м.п.м.

Изучим оптимальные свойства точности метода м.п.м. в сравнении с другими методами на классах систем с приближенными данными.

Пусть для простоты $\mathfrak{A}_0 = \mathfrak{A}$. Введем следующее

Определение 1. Назовем *стандартным методом приближенного нахождения нормального псевдорешения СЛАУ* (1.1) всякое отображение P, ставящее в соответствие каждому набору приближенных данных $(A_h, u_\sigma, h, \sigma)$, где $0 \leqslant h < H_0$, $0 \leqslant \sigma \leqslant \sigma_0$, задачи (1.1) элемент $z_\eta = P(A_h, u_\sigma, h, \sigma) \in \mathbf{R}^n$, который может быть представлен в виде $z_\eta = V_h F(D_h, \alpha_\eta) U_h^T u_\sigma$. Здесь $\alpha_\eta \geqslant 0$ — параметр регуляризации, выбираемый некоторым образом и зависящий в общем случае от $(A_h, u_\sigma, h, \sigma)$; матрица $F(D_h, \alpha_\eta) \in \mathfrak{A}^*$ — матрица вида

$$F(D_h, \alpha_\eta) = \operatorname{diag}[f_1(\rho_1^h, \alpha_\eta), \ldots, f_M(\rho_M^h, \alpha_\eta)];$$

входящие в нее функции $f_k(\rho, \alpha)$ $(k = 1, \ldots, M)$ определены при ρ, $\alpha \geqslant 0$, измеримы по ρ и ограничены на полуинтервалах вида $\rho \geqslant \varepsilon$ (для любого $\varepsilon > 0$) при каждом фиксированном $\alpha \geqslant 0$.

Определение допускает случай, когда зависимость от α_η отсутствует. Так, например, для метода наименьших квадратов $z_\eta = A_h^+ u_\sigma = V_h D_h^+ U_h^T u_\sigma$, $f_k = \theta(\rho_k^h)$ $(k = 1, \ldots, M)$. К стандартным методам относятся также такие алгоритмы решения СЛАУ, как метод регуляризации:

$$z_\eta = (\alpha_\eta I + A_h^T A_h)^{-1} A_h^T u_\sigma = V_h (\alpha_\eta I + D_h^T D_h)^{-1} D_h^T U_h^T u_\sigma,$$

$$f_k = [\alpha_\eta + (\rho_k^h)^2]^{-1} \rho_k^h;$$

метод, изучавшийся в работах [28, 55]:

$$z_\eta = (\alpha_\eta I + A_h^T A_h)^{-1} A_h^T A_h A_h^T (\alpha_\eta E + A_h A_h^T)^{-1} u_\sigma =$$

$$= V_h (\alpha_\eta I + D_h^T D_h)^{-1} D_h^T D_h D_h^T (\alpha_\eta E + D_h D_h^T) U_h^T u_\sigma,$$

$$f_k = [\alpha_\eta + (\rho_k^h)^2]^{-2} (\rho_k^h)^3$$

и др. (см., например, [25, 143, 232]).

Теорема 1. *Метод м.п.м. для решения основной задачи* § 1 *является стандартным при* $0 \leqslant h < H_0(\bar{A}) = \mathrm{const}$.

Доказательство. Как указано в § 1, $\bar{A} \neq 0$. Поэтому найдется такое число $H_0(\bar{A}) = \min(\|\bar{A}^+\|_*^{-1}/2, \|\bar{A}\|/2)$, что при $0 \leqslant h < H_0(\bar{A})$ будут выполнены условия теорем 3.6, 3.2. По этим теоремам для приближенного решения z_η основной задачи, получаемого по методу м.п.м, справедливо представление

$$z_\eta = \hat{A}_h^+ u_\sigma = V_h \hat{D}_h^+ U_h^T u_\sigma, \quad \hat{D}_h^+ = \operatorname{diag}[\theta(\hat{\rho}_1), \ldots, \theta(\hat{\rho}_M)].$$

Укажем явный вид функций $f_k(\rho, \alpha)$ из определения 1 для метода м.п.м.:

$$f_k=f(\rho, \alpha)=\{[\rho X(\alpha, \rho)]^{-1}\colon \rho>0,\ 0\leqslant\alpha<27\rho^4/16;\ 0\colon \rho\geqslant 0,\ \alpha\geqslant 27\rho^4/16\}.$$

Здесь $X(\alpha, \rho)$ определяется как положительное решение уравнения $X^4-X^3=\alpha\rho^{-4}$ (ср. (3.10′)). Согласно формуле (3.9), $f(\rho_k^h, \lambda)=\theta[\rho_k(\lambda)]$, а по теореме 3.2 $\theta(\hat{\rho}_k)=\theta\{\rho_k[\lambda(h)]\}=f[\rho_k^h, \lambda(h)]$. Фигурирующее здесь число $\lambda(h)>0$ однозначно находится из уравнения (3.14) и служит параметром регуляризации: $\alpha_\eta=\lambda(h)$. Для завершения доказательства заметим, что функция $f(\rho, \alpha)$ удовлетворяет требованиям определения 1, так как величина $X(\alpha, \rho)$ непрерывно зависит от α, ρ в совокупности при $\alpha\geqslant 0$, $\rho>0$ (см. свойство 4 решения уравнения (3.10′)).

Будем обозначать множество всех стандартных методов как S. Определим при фиксированных числах R_1, $R_2>0$ класс $\Sigma\equiv\Sigma(R_1, R_2)=\{(\bar{A}, \bar{u})\colon \|\bar{A}^+\|_*\leqslant R_1,\ \|\bar{u}\|\leqslant R_2\}$ точных систем вида (1), а для каждой точной системы $(\bar{A}, \bar{u})\in\Sigma$ зададим класс приближенных данных $\Sigma_\eta\equiv\Sigma_\eta(\bar{A}, \bar{u})=\{(A_h, u_\sigma)\colon \|A_h-\bar{A}\|\leqslant h,\ \|u_\sigma-\bar{u}\|\leqslant\sigma\}$, отвечающих точностям $\eta=(h, \sigma)$. Введем характеристику точности метода $P\in S$ на классе $\Sigma(R_1, R_2)$ точных систем, если приближенные данные берутся из соответствующих классов $\Sigma_\eta(\bar{A}, \bar{u})$:

$$\Delta(P, \eta)\equiv\Delta(P, \eta, R_1, R_2)=$$

$$=\sup_{A_h,\, u_\sigma,\, \bar{A},\, \bar{u}}\{\|P(A_h, u_\sigma, h, \sigma)-\bar{z}\|\colon \bar{z}=\bar{A}^+\bar{u},$$

$$(\bar{A}, \bar{u})\in\Sigma(R_1, R_2),\ (A_h, u_\sigma)\in\Sigma_\eta(\bar{A}, \bar{u})\}. \qquad (1)$$

Определение 2. *Оптимальной точностью* приближенного решения систем (1) из $\Sigma(R_1, R_2)$ на классе стандартных методов назовем величину

$$\Delta_0(\eta)\equiv\Delta_0(\eta, R_1, R_2)=\inf\{\Delta(P, \eta)\colon P\in S\}.$$

Определение 3. Метод $P\in S$ имеет *оптимальный порядок точности*, если

$$\overline{\lim_{\eta\to 0}}[\Delta(P, \eta, R_1, R_2)/\Delta_0(\eta, R_1, R_2)]\leqslant g,$$

где константа $g\geqslant 1$ не зависит от R_1, R_2.

Получим нижнюю оценку для точности произвольного метода $P\in S$.

Лемма 1. *Для любого $P\in S$ и всякого* $\eta=(h, \sigma)$ $(0\leqslant h\leqslant h_0(R_1)\equiv\equiv R_1^{-1}/2,\ 0\leqslant\sigma<R_2)$ *выполнена оценка*

$$\Delta(P, \eta)\geqslant R_1(\sigma+R_1R_2h)[1-2hR_1/q(h)](1+q(h))^{-2}(1+\sigma/R_2)^{-1}\equiv\nu(h, \sigma),$$

где $q(h)$ — произвольная ограниченная функция, подчиненная неравенству $q(h)>2hR_1$.

Доказательство. Выберем из множества $\Sigma(R_1, R_2)$ подкласс Σ_1 точных систем с данными $(\bar{A}, \bar{u})$, удовлетворяющими условиям $\bar{A}=\bar{U}\bar{D}\bar{V}^T$, где $\bar{U}\in\mathfrak{A}_m$, $\bar{V}\in\mathfrak{A}_n$ — ортогональные матрицы, а $\bar{D}=\mathrm{diag}(\bar{\rho}, 0, \dots, 0)\in\mathfrak{A}$ $(\bar{\rho}>0)$; $\bar{u}\in\mathrm{Im}\,\bar{A}$, $\|\bar{u}\|=R_2$. Из соответствующего каждому такому набору $(\bar{A}, \bar{u})$ множества $\Sigma_\eta(\bar{A}, \bar{u})$ выделим подмножество Σ_η^1 приближенных данных (A_h, u_σ), которые удовлетворяют требованиям

$$A_h = \bar{U} D_h \bar{V}^T, \quad D_h = \operatorname{diag}(\tilde{\rho}, 0, \ldots, 0) \in \mathfrak{A},$$
$$|\bar{\rho} - \tilde{\rho}| \leqslant h; \quad u_\sigma = t\bar{u}, \quad |t-1| \leqslant \sigma/R_2.$$

Тогда на множествах Σ_1, Σ_η^1 очевидны соотношения

$$\|\bar{A} - A_h\| = \|\bar{D} - D_h\| = |\bar{\rho} - \tilde{\rho}| \leqslant h; \quad \|\bar{A}^+\|_* = \bar{\rho}^{-1} \leqslant R_1;$$
$$\|\bar{u} - u_\sigma\| = |t-1| \cdot \|\bar{u}\| \leqslant \sigma.$$

Используя определение стандартного метода и учитывая принятые предположения о $\bar{A}$, $\bar{u}$, A_h, u_σ, получаем

$$\|P(A_h, u_\sigma, h, \sigma) - \bar{A}^+\bar{u}\|^2 = \|\bar{V}F(D_h, \alpha_\eta)\bar{U}^T(t\bar{u}) - \bar{V}\bar{D}^+\bar{U}^T\bar{u}\|^2 =$$
$$= \sum_{k=1}^{M} \bar{u}_k^2 [f_k(\rho_k^h, \alpha_\eta)t - \theta(\bar{\rho}_k)]^2 = \bar{u}_1^2 [f_1(\tilde{\rho}, \alpha_\eta)t - \bar{\rho}^{-1}]^2, \tag{2}$$

где $\bar{u}_k$ $(k=1, \ldots, M)$ — координаты вектора $\bar{U}^T\bar{u} \in \mathbf{R}^m$. При получении этого равенства учтено также, что в силу принадлежности $\bar{u} \in \operatorname{Im} \bar{A}$ вектор $\bar{U}^T\bar{u}$ имеет вид $\bar{U}^T\bar{u} = (\bar{u}_1, 0, \ldots, 0)$, $|\bar{u}_1| = \|\bar{u}\| = R_2$.

Из (1), (2) поэтому следует

$$\Delta^2(P, \eta) = \sup_{\bar{A}, \bar{u}, A_h, u_\sigma} \{\|P(A_h, u_\sigma, h, \sigma) - \bar{A}^+\bar{u}\|^2 : (\bar{A}, \bar{u}) \in \Sigma(R_1, R_2),$$
$$(A_h, u_\sigma) \in \Sigma_\eta(\bar{A}, \bar{u})\} \geqslant \sup_{\bar{A}, \bar{u}, A_h, u_\sigma} \{\|P(A_h, u_\sigma, h, \sigma) - \bar{A}^+\bar{u}\|^2 : (\bar{A}, \bar{u}) \in \Sigma_1,$$
$$(A_h, u_\sigma) \in \Sigma_\eta^1\} = \sup_{\bar{\rho}, \bar{u}_1, \tilde{\rho}, t} \{\bar{u}_1^2 [f_1(\tilde{\rho}, \alpha_\eta)t - \bar{\rho}^{-1}]^2 : |\bar{\rho} - \tilde{\rho}| \leqslant h, \ \bar{\rho}^{-1} \leqslant R_1; \ |\bar{u}_1| =$$
$$= R_2; \ |t-1| \leqslant \sigma/R_2\} \geqslant$$
$$\geqslant R_2^2 \sup_{\bar{\rho}} \{[f_1(\bar{\rho} \pm h, \alpha_0)(1 \mp \sigma/R_2) - \bar{\rho}^{-1}]^2 : \bar{\rho} \geqslant R_1^{-1}\} \equiv \Delta_\pm^2.$$

В этой выкладке число α_0 — фиксированное значение параметра регуляризации, взятое из области значений функции $\alpha_\eta = \alpha_\eta(A_h, u_\sigma, h, \sigma)$. Величины $\Delta_\pm$ конечны, так как при $\bar{\rho} \geqslant R_1^{-1}$ и при $0 \leqslant h < h_0(R_1) = R_1^{-1}/2$ оказывается, что $\bar{\rho} - h > R_1^{-1}/2 > 0$, и, согласно свойствам функции f_1 (см. определение 1), множество значений $f_1(\bar{\rho} - h, \alpha_0)$ ограничено.

Установленное неравенство для $\Delta^2(P, \eta)$ позволяет получить оценку

$$\Delta(P, \eta) \geqslant (\Delta_+ + \Delta_-)/2 \geqslant R_2 \{|f_1(\rho + h, \alpha_0)(1 - \sigma/R_2) - \rho^{-1}| +$$
$$+ |f_1(\rho - h, \alpha_0)(1 + \sigma/R_2) - \rho^{-1}|\}/2, \tag{3}$$

которая выполняется при $\rho \geqslant R_1^{-1}$ и при $0 \leqslant h < h_0(R_1)$. Обозначим $\beta \equiv R_1^{-1}$, $\gamma \equiv [1 + q(h)] R_1^{-1}$. Тогда вследствие условия на функцию

$q(h)$ обеспечивается неравенство $\gamma-h>\beta+h$. С учетом этого из (3) следует

$$\Delta(P,\ \eta)\geqslant\frac{R_2}{2(\gamma-\beta)}\int\limits_{\beta}^{\gamma}\{|f_1(\rho+h,\ \alpha_0)\left(1-\frac{\sigma}{R_2}\right)-\rho^{-1}|+$$

$$+|f_1(\rho-h,\ \alpha_0)\left(1+\frac{\sigma}{R_2}\right)-\rho^{-1}|\}\,d\rho\geqslant$$

$$\geqslant R_2\frac{1}{2(\gamma-\beta)}\int\limits_{\beta+h}^{\gamma-h}\{|f_1(\rho,\ \alpha_0)\left(1-\frac{\sigma}{R_2}\right)-(\rho-h)^{-1}|+$$

$$+|f_1(\rho,\ \alpha_0)\left(1+\frac{\sigma}{R_2}\right)-(\rho+h)^{-1}|\}\,d\rho\geqslant$$

$$\geqslant\frac{R_2(\gamma-\beta-2h)}{2(\gamma-\beta)}\left[2\left(h+\frac{\sigma\beta}{R_2}\right)\gamma^{-2}\left(1+\frac{\sigma}{R_2}\right)^{-1}\right]=$$

$$=\left[1-\frac{2hR_1}{q(h)}\right](\sigma+R_1R_2h)\,R_1\,[1+q(h)]^{-2}\left(1+\frac{\sigma}{R_2}\right)^{-1}.$$

Здесь кроме замены переменного в интеграле использовано очевидное неравенство

$$|ax-c|+|bx-d|\geqslant|bc-ad|\,/\max\{|a|,\ |b|\}$$

с $x=f_1(\rho,\ \alpha_0)$, $a,\ b=(1\mp\sigma/R_2)>0$, $c,\ d=(\rho\mp h)^{-1}$. Лемма доказана.

Из определения 2 и леммы 1 легко получить

Следствие 1. *При* $0\leqslant h<h_0(R_1)$ *выполнена оценка* $\Delta_0(\eta)\geqslant\nu(h,\ \sigma)$.

Введем теперь в соответствии с формулой (1) величину $\Delta_{\text{м.п.м.}}(\eta)$, характеризующую точность метода м.п.м.

Лемма 2. *Характеристика* (1) *точности метода м.п.м. имеет следующий порядок при* h, $\sigma\to 0$:

$$\Delta_{\text{м.п.м.}}(\eta)=R_1(\sigma+2hR_1R_2)[1+o(h+\sigma)].$$

В этой оценке порядок входящих в нее величин h, σ, R_1, R_2, *а также константа* 2 *точные.*

Доказательство. Используем тот факт, что данные $(\bar{A},\ \bar{u})$ входят в Σ, т. е. что $\|\bar{A}^+\|_*\leqslant R_1$, $\|\bar{u}\|\leqslant R_2$. Тогда $h_0(R_1)=$ $=R_1^{-1}/2\leqslant\|\bar{A}^+\|_*^{-1}/2=h_0(\bar{A})$ и выполнение неравенства $0\leqslant h<h_0(R_1)$ влечет за собой выполнение условий теоремы 2.6. Из этой теоремы следует оценка (2.16), которую запишем в виде

$$\|z_\eta(A_h,\ u_\sigma,\ h,\ \sigma)-\bar{z}\|\leqslant\|\bar{A}^+\|_*[\sigma+2h\|\bar{A}^+\|_*\cdot\|\bar{u}\|/(1-2h\|\bar{A}^+\|_*)^3],$$

где z_η — приближенное решение основной задачи по методу м.п.м. Из этой оценки и определения (1) точности метода м.п.м. получим

$$\Delta_{\text{м.п.м.}}(\eta)\leqslant R_1[\sigma+2hR_1R_2/(1-2hR_1)^3]. \tag{4}$$

Из определения (1) ясно также, что $\Delta_{\text{м.п.м.}}(\eta) \geq \| z_\eta(A_h, u_\sigma, h, \sigma) - \bar{z} \|$ для любой точной задачи (1) с данными $(\bar{A}, \bar{u})$ из Σ и любых приближенных данных $(A_h, u_\sigma) \in \Sigma_\eta(\bar{A}, \bar{u})$. Выберем конкретные точные данные $(\bar{A}, \bar{u})$:

$$\bar{A} = \mathrm{diag}(R_1^{-1}, 0, \ldots, 0) \in \mathfrak{A}, \quad \bar{u} = (R_2, 0, \ldots, 0) \in \mathbf{R}^m,$$

а также определенные приближенные данные

$$A_h = \mathrm{diag}(R_1^{-1} + h, 0, \ldots, 0) \in \mathfrak{A}, \quad u_\sigma = (R_2 - \sigma, 0, \ldots, 0) \in \mathbf{R}^m.$$

Используя методику из § 3, можно вычислить матрицу метода м.п.м.: $\hat{A}_h = \mathrm{diag}\,[(R_1^{-1} + h)\, x\, [\lambda(h)],\ 0,\ \ldots,\ 0]$. Здесь использована теорема 3.2. При этом $x\,[\lambda(h)]$ находится из уравнения типа (3.14):

$$\beta(\lambda) = (R_1^{-1} + h)^2 [x(\lambda) - 1]^2 = h^2,$$

откуда $x\,[\lambda(h)] = 1 + h/(R_1^{-1} + h)$. Таким образом, $\hat{A}_h = \mathrm{diag}(R_1^{-1} + 2h, 0, \ldots, 0)$. Для рассматриваемых данных поэтому

$$\bar{z} = \bar{A}^+ \bar{u} = (R_1 R_2, 0, \ldots, 0) \in \mathbf{R}^n,$$
$$z_\eta = \hat{A}_h^+ u_\sigma = ([R_2 - \sigma] / [R_1^{-1} + 2h], 0, \ldots, 0) \in \mathbf{R}^n.$$

Отсюда приходим к оценке

$$\Delta_{\text{м.п.м.}}(\eta) \geq \| z_\eta - \bar{z} \| = (\sigma R_1 + 2h R_1^2 R_2) / (1 + 2h R_1). \tag{5}$$

Сопоставляя оценки (4) и (5), можно найти гланый член величины $\Delta_{\text{м.п.м.}}(\eta)$ при $\eta \to 0$: $\Delta_{\text{м.п.м.}}(\eta) \asymp R_1(\sigma + 2h R_1 R_2)$. Лемма доказана.

Из лемм 1, 2 и следствия 1 получается

Теорема 2. *Метод м.п.м. как алгоритм класса S имеет оптимальный порядок точности*:

$$\overline{\lim_{\eta \to 0}} \left[\Delta_{\text{м.п.м.}}(\eta, R_1, R_2) / \Delta_0(\eta, R_1, R_2) \right] \leq 2.$$

Действительно, из следствия 1 и из (4) имеем

$$\nu(h, \sigma) \leq \Delta_0(\eta) \leq \Delta_{\text{м.п.м.}} \leq R_1 [\sigma + 2h \cdot R_1 R_2 / (1 - 2h R_1)^3]. \tag{6}$$

Выберем теперь функцию $q(h)$ из леммы 1 в виде $q(h) = 2\sqrt{h} R_1$. Тогда, учитывая вид величины $\nu(h, \sigma)$, указанный в лемме 1, получим из (6) при $0 \leq h < \min\{1, h_0(R_1)\}$

$$1 \leq \Delta_{\text{м.п.м.}}(\eta) / \Delta_0(\eta) \leq 2(1 + 2\sqrt{h} R_1)^2 (1 + \sigma/R_2)(1 - 2h R_1)^{-3} (1 - \sqrt{h})^{-1}.$$

Переходя в этом неравенстве к пределу при $h, \sigma \to 0$, получаем доказываемое в данной теореме утверждение.

Отметим, что из (6) следует сходимость $\Delta_0(\eta) \to 0$ при $\eta \to 0$.

Итак, метод м.п.м. имеет оптимальный порядок точности с константой оптимальности $g = 2$. Можно сравнить этот метод с другими, вычисляя для них аналогичную константу. Тогда, например, для метода из работ [28, 55] оказывается $g = (7 + 2\sqrt{5})^{1/2} \approx 3{,}387$.

Метод м.п.м. обладает также другим свойством оптимальности — *свойством оптимального порядка близости числа обусловленности матрицы метода м.п.м. к наилучшему.* Чтобы рассмотреть это свойство, введем число обусловленности матрицы $A: \nu(A)=\|A^+\|_* \cdot \|A\|$. Известно, что число $\nu(A)$ характеризует устойчивость задачи нахождения нормального псевдорешения СЛАУ (1.1) (см. [45, 46]). Поэтому при практических расчетах предпочтительнее использовать приближенную матрицу A_h с возможно меньшим числом обусловленности. В связи с этим введем экстремальную задачу: найти матрицу $\breve{A}_h \in \mathfrak{A}_h$, для которой

$$\nu(\breve{A}_h)=\inf\{\nu(A): A\in\mathfrak{A}_h\}=\inf\{\nu(A): A\in\mathfrak{A},\ \|A-A_h\|\leqslant h\}\equiv\nu_h. \quad (7)$$

Будем называть число ν_h *наилучшим возможным числом обусловленности* задачи (1.1) с приближенной матрицей A_h.

Теорема 3. *Если при заданном* $h\geqslant 0$ *выполнено условие* $\|A_h\|>h$, *то задача* (7) *разрешима.*

Доказательство. Пусть $\{A_n\}$ — произвольная минимизирующая последовательность для задачи (7), т. е. $\{A_n\}\subset\mathfrak{A}_h$, $\nu(A_n)\to\nu_h$ при $n\to\infty$. Тогда при $n\geqslant N_0$ существует такое число $\varepsilon(N_0)>0$, что

$$\nu(A_n)=\|A_n^+\|_* \cdot \|A_n\|\leqslant\nu_h+\varepsilon(N_0)\equiv M_0=\text{const}. \quad (8)$$

Множество $\{A_n\}$ компактно в $\mathfrak{A}$ в силу своей ограниченности: $\|A_n\|\leqslant\|A_h\|+\|A_n-A_h\|\leqslant\|A_h\|+h$. Оно также замкнуто в $\mathfrak{A}$ вследствие вложения $\{A_n\}\subset\mathfrak{A}_h$ (см. § 2). Поэтому существует подпоследовательность $\{A_{n_k}\}$, сходящаяся при $k\to\infty$ к $A_0\in\mathfrak{A}_h$. При этом $A_0\neq 0$. Действительно, если бы было выполнено равенство $A_0=0$, то выполнялось бы противоречивое соотношение

$$h\geqslant\lim_{k\to\infty}\|A_{n_k}-A_h\|=\|A_0-A_h\|=\|A_h\|>h.$$

Из неравенства $A_0\neq 0$ вытекает оценка $\|A_{n_k}\|\geqslant q=\text{const}>0$ при $n_k\geqslant N_1$. Тогда при $n_k>\max(N_0, N_1)$ получим из (8) $\|A_{n_k}^+\|_*\leqslant M_0/q\equiv M_1$ и, кроме того, $\|A_{n_k}\|\leqslant\|A_h\|+h\equiv M_2$. Из этих оценок по лемме 2.1 приходим к выводу, что $A_{n_k}^+\to A_0^+$ при $k\to\infty$. Поэтому

$$\lim_{k\to\infty}\nu(A_{n_k})=\lim_{k\to\infty}\|A_{n_k}^+\|_*\cdot\lim_{k\to\infty}\|A_{n_k}\|=\|A_0^+\|_*\cdot\|A_0\|=\nu(A_0)=\nu_h,$$

т. е. A_0 есть решение задачи (7). Теорема доказана.

Назовем матрицу $\breve{A}_h$ — любое решение задачи (7) — *оптимальной по числу обусловленности на множестве* $\mathfrak{A}_h$. Отметим при этом, что множество $\mathfrak{A}_h$ содержит элементы $\bar{A}$, A_h. Условие теоремы 3 выполнено при $0\leqslant h<\|\bar{A}\|/2$.

Для исследования свойств матрицы $\breve{A}_h$ потребуется

Лемма 3. *Пусть* p, q — *фиксированные положительные числа и* $\bar{h}\equiv\min(1, p^{-1})q^{-1}\|\bar{A}^+\|_*^{-1}$. *Если при* $0\leqslant h<\bar{h}$ *для матрицы* A *выполнены неравенства* $\|A-\bar{A}\|\leqslant qh$, $\|A^+\|_*\leqslant p\|\bar{A}^+\|_*$, *то* $\operatorname{Rg} A=\operatorname{Rg}\bar{A}$.

Доказательство этой леммы проводится по той же схеме, что и в теореме 2.4.

Теорема 4. *При $h\to 0$ имеет место сходимость $\breve{A}_h^+\to\bar{A}^+$. Если $\breve{z}_\eta\equiv\breve{A}_h^+u_\sigma$, то $\breve{z}_\eta\to\bar{z}$ при $\eta\to 0$.*

Доказательство. Пусть h таково, что $0\leqslant h<\|\bar{A}\|/4$. Тогда по теореме 3 решение задачи (7) существует, и для него выполнены оценки

$$\|\breve{A}_h-\bar{A}\|\leqslant\|\breve{A}_h-A_h\|+\|A_h-\bar{A}\|\leqslant 2h,$$
$$\|\breve{A}_h^+\|_*=\nu(\breve{A}_h)\|\breve{A}_h\|^{-1}\leqslant\nu(\bar{A})/(\|\bar{A}\|-2h)\leqslant 2\|\bar{A}^+\|_*.$$

Тогда по лемме 3 при $0\leqslant h<\hat{h}\equiv\min\{\|\bar{A}\|/4,\ \|\bar{A}^+\|_*^{-1}/4\}$ ранги матриц $\breve{A}_h$ и $\bar{A}$ совпадают. В силу этого по лемме 2.3 справедливо неравенство

$$\|\breve{A}_h^+-\bar{A}^+\|_*\leqslant\|\breve{A}_h-\bar{A}\|\cdot\|\bar{A}^+\|_*^2/(1-\|\breve{A}_h-\bar{A}\|\,\|\bar{A}^+\|_*)^3\leqslant$$
$$\leqslant 2h\,\|\bar{A}^+\|_*^2/(1-2h\,\|\bar{A}^+\|_*)^3,$$

из которого вытекает искомая сходимость $\breve{A}_h^+\to\bar{A}^+$ при $h\to 0$. Из оценки типа (2.6) для величины $\|\breve{z}_\eta-\bar{z}\|$ следует тогда сходимость $\breve{z}_\eta\to\bar{z}$ при $\eta\to 0$.

Теорема 4 обосновывает устойчивость алгоритма приближенного нахождения нормального псевдорешения СЛАУ (1.1), основанного на использовании приближений $\breve{z}_\eta$. Из этой теоремы следует также, что $\nu(\breve{A}_h)\to\nu(\bar{A})$ при $h\to 0$.

Хотя матрица $\breve{A}_h$ и обладает важным свойством минимальности числа обусловленности, реализация процедуры нахождения таких матриц встречает определенные трудности. В связи с этим имеет смысл рассматривать алгоритмы устойчивого определения $\bar{z}$, использующие почти оптимальные по числу обусловленности матрицы. К их числу относится и метод м.п.м.

Определение 4. Матрица $A(h)\in\mathfrak{A}_h$ имеет *оптимальный порядок близости к наилучшему числу обусловленности*, если

$$\nu[A(h)]-\nu(\breve{A}_h)\leqslant lh\,\|\bar{A}^+\|_*,$$

где константа $l\geqslant 1$ не зависит от h, $\bar{A}$.

Теорема 5. *Матрица метода м.п.м. имеет оптимальный порядок близости к наилучшему числу обусловленности.*

Доказательство. Возьмем любую матрицу $\tilde{A}_h$ метода м.п.м. (см. § 2) и произвольное решение $\breve{A}_h$ задачи (7). Для них справедливы оценки

$$\|\tilde{A}_h^+\|_*\leqslant\|\bar{A}^+\|_*,\qquad\|\tilde{A}_h^+\|_*\leqslant\|\breve{A}_h^+\|_*,$$

$$\|\breve{A}_h\|\geqslant\|\tilde{A}_h\|-\|\breve{A}_h-\tilde{A}_h\|\geqslant\|\tilde{A}_h\|-2h.$$

Используя их, можно получить неравенство

$$\nu(\breve{A}_h)=\|\breve{A}_h^+\|_*\cdot\|\breve{A}_h\|\geqslant\|\tilde{A}_h^+\|_*(\|\tilde{A}_h\|-2h)\geqslant\nu(\tilde{A}_h)-2h\,\|\bar{A}^+\|_*,$$

из которого следует, что

$$\nu(\tilde{A}_h)-\nu(\breve{A}_h)\leqslant 2h\,\|\bar{A}^+\|_*.$$

Это и доказывает теорему.

§ 6. О численной реализации метода м.п.м.

Будем рассматривать этот вопрос в случае, когда $\mathfrak{A}_0=\mathfrak{A}$ и можно получить матрицу метода м.п.м. на основании теории, развитой в § 3. Схематически *алгоритм, основанный на методе м.п.м.* состоит из следующих частей [131].

1. Нахождение сингулярного разложения заданной матрицы A_h: $A_h=U_hD_hV_h^T$.

2. Нахождение решения $\lambda(h)$ уравнения (3.14).

3. Построение матрицы метода $\hat{A}_h=U_h\hat{D}_hV_h^T$, где $\hat{D}_h$ определяется на основании теоремы 3.2 или задается, как в подходе, базирующемся на теореме 3.4.

4. Определение приближенного решения $z_\eta=\hat{A}_h^+u_\sigma=V_h\hat{D}_h^+U_h^Tu_\sigma$. Вычисление при необходимости самой минимальной псевдообратной матрицы $\hat{A}_h^+=V_h\hat{D}_h^+U_h^T$.

Рассмотрим более подробно каждую часть алгоритма. В настоящее время имеется ряд методов приближенного нахождения сингулярного разложения матриц общего вида [44, 143, 216]. Опубликованы и используются на практике соответствующие программы для ЭВМ на различных языках высокого уровня [167, 212, 216]. Эти методы дают, вообще говоря, *почти сингулярное разложение с точностью* $\varkappa$ заданной матрицы A_h. В результате их применения получаются ортогональные матрицы $U_{h\varkappa}\in\mathfrak{A}_m$, $V_{h\varkappa}\in\mathfrak{A}_n$, а также диагональная матрица $D_{h\varkappa}=\mathrm{diag}(\rho_1^{h\varkappa},\ldots,\rho_M^{h\varkappa})\in\mathfrak{A}$ $(\rho_1^{h\varkappa}\geqslant\ldots\geqslant\rho_M^{h\varkappa}\geqslant 0)$, причем

$$\|A_h-U_{h\varkappa}D_{h\varkappa}V_{h\varkappa}^T\|\leqslant\varkappa. \tag{1}$$

Поскольку $\varkappa$ — обычно достаточно малая величина ($\varkappa\ll h$), то почти сингулярное разложение можно использовать в алгоритме вместо точного. При этом следует считать, что задана не матрица A_h, а матрица $A_{h\varkappa}=U_{h\varkappa}D_{h\varkappa}V_{h\varkappa}^T$, для которой очевидно выполнено условие аппроксимации $\|\bar{A}-A_{h\varkappa}\|\leqslant h+\varkappa$. Тем самым вместо данных (A_h, u_σ, h) задачи (1.1), используемых в методе м.п.м., можно использовать другие данные ($A_{h\varkappa}$, u_σ, $h+\varkappa$). Тогда в расчетных формулах алгоритма, приведенных в § 3, вместо величин U_h, D_h, V_h, h следует поставить $U_{h\varkappa}$, $D_{h\varkappa}$, $V_{h\varkappa}$, $h+\varkappa$.

Такой подход снимает ряд проблем. Действительно, нахождение точного сингулярного разложения матрицы A_h эквивалентно решению полной проблемы собственных значений для матриц $A_h^TA_h$, $A_hA_h^T$ (см. [46]). При этом возникают вопросы точности и устойчивости определения собственных векторов и собственных значений. Алгоритмы нахождения почти сингулярного разложения этих вопросов не касаются, так как в них достаточно найти любые матрицы $U_{h\varkappa}$, $D_{h\varkappa}$, $V_{h\varkappa}$, для которых будут выполнены указанные выше условия, в частности (1). Тогда столбцы матриц $U_{h\varkappa}$, $V_{h\varkappa}$ уже, вообще говоря, не будут собственными векторами матриц $A_h^TA_h$, $A_hA_h^T$. Однако для метода м.п.м. это не существенно.

Часть 2 алгоритма связана с решением уравнения (3.14) с монотонной функцией

$$\beta(\lambda)=\sum_{k=1}^{M}[\rho_k(\lambda)-\rho_k^{h\varkappa}]^2=(h+\varkappa)^2. \tag{2}$$

Левая часть этого уравнения записывается на основании леммы 3.2 следующим образом:

$$\beta(\lambda)=\sum_{k\in K_1(\lambda)}(\rho_k^{h\varkappa})^2\left[x_k(\lambda)-1\right]^2+\sum_{k\in K\setminus K_1(\lambda)}(\rho_k^{h\varkappa})^2,\ \ \lambda>0. \tag{3}$$

В этой формуле множества индексов K, $K_1(\lambda)$ определяются так: $K=\{k=1, \ldots, M\}$, $K_1\equiv K_1(\lambda)\equiv K_1(\lambda,\ A_{h\varkappa})=\{k\in K\colon \lambda(\rho_k^{h\varkappa})^{-4}\leqslant 27/16\}$. Величина $x_k(\lambda)$ представляет собой положительный корень уравнения

$$x^4-x^3=\lambda(\rho_k^{h\varkappa})^{-4}, \tag{4}$$

лежащий на $(1, 3/2]$.

Нахождение решения $\lambda(h+\varkappa)$ уравнения (2) существенно использует процедуру вычисления функции $\beta(\lambda)$. Для такого вычисления необходимо при заданном λ определить по сингулярным числам $\rho_k^{h\varkappa}$ матрицы $A_{h\varkappa}$ множество индексов $K_1(\lambda)$, найти для $k\in K_1(\lambda)$ решение $x_k(\lambda)$ уравнения (4) и произвести расчет функции $\beta(\lambda)$ по формуле (3). Анализируя значение функции $\beta(\lambda)$ в ее точках разрыва $\lambda_k=27(\rho_k^{h\varkappa})^4/16$ $(k\in K)$ (см. лемму 3.3) и сравнивая их с $(h+\varkappa)^2$, можно выяснить, обычное решение имеет уравнение (2) или обобщенное. При этом нахождение обобщенного решения, совпадающего с одной из точек разрыва, требует конечного числа вычислений значений $\beta(\lambda)$. Сама процедура вычисления $\beta(\lambda)$ проводится за относительно малое число операций, что позволяет использовать для поиска обычного решения уравнения (2), например, метод дихотомии.

Реализация части 3 алгоритма зависит от того, будет ли решение $\lambda(h+\varkappa)$ уравнения (2) обычным или обобщенным. Если это решение обычное, то по теореме 3.2 матрица $\hat{D}_{h\varkappa}$, содержащая сингулярные числа матрицы метода м.п.м., имеет вид

$$\hat{D}_{h\varkappa}=\operatorname{diag}\{\rho_1[\lambda(h+\varkappa)], \ldots, \rho_M[\lambda(h+\varkappa)]\}\equiv\operatorname{diag}(\rho_1^*, \ldots, \rho_M^*)\in\mathfrak{A},$$

где зависимости $\rho_k(\lambda)$ $(k\in K)$ определяются формулой (3.9). Согласно последней, $\rho_k^*=\{\rho_k^{h\varkappa}x_k[\lambda(h+\varkappa)]\colon k\in K_1[\lambda(h+\varkappa)];\ 0\colon k\in K\setminus K_1[\lambda(h+\varkappa)]\}$.

В случае, если $\lambda(h+\varkappa)$ — обобщенное решение уравнения (2), т. е. это число совпадает с одной из точек разрыва функции $\beta(\lambda)$, то можно использовать подход, обоснованный в теореме 3.4. При этом, если считать, что $\lambda(h+\varkappa)=\lambda_{t(h+\varkappa)+1}$, то можно положить $\hat{D}_{h\varkappa}=\operatorname{diag}(\rho_1^*, \ldots, \rho_M^*)$. Здесь числа ρ_k^* $(k\in K)$ определяются, как в § 3:

$$\rho_k^*=\{\rho_k^{h\varkappa}x_k[\lambda_{t(h+\varkappa)+1}]\colon k=1, \ldots, t(h+\varkappa);\ 0\colon k=t(h+\varkappa)+1, \ldots, M\}.$$

При $0<h+\varkappa<h_1(\bar{A})$ по теореме 3.3 справедливо равенство $t(h+\varkappa)=$ $=\operatorname{Rg}\bar{A}$.

Заключая описание алгоритма, отметим, что для вычисления приближенного решения $z_{\eta\varkappa}=V_{h\varkappa}\hat{D}_{h\varkappa}^{+}U_{h\varkappa}^{T}u_\sigma$ используется псевдообратная матрица $\hat{D}_{h\varkappa}^{+}=\operatorname{diag}[\theta(\rho_1^*), \ldots, \theta(\rho_M^*)]\in\mathfrak{A}^*$.

На основании теорем сходимости 2.7, 3.4 можно утверждать, что $z_{\eta\varkappa}\to\bar{z}$ при η, $\varkappa\to 0$.

Отметим, что в случае, когда $\lambda(h+\varkappa)$ — обобщенное решение уравнения (2), можно построить матрицу метода м.п.м. на основе теоремы 3.5 (а также леммы 3.4 и теоремы 3.1). Для этого можно использовать устойчивую процедуру вычисления ранга матрицы $\bar{A}$, указанную в теореме 3.3.

Можно оценить число N арифметических операций, необходимых для нахождения приближенного решения $z_{\eta\varkappa}$ в случае квадратной матрицы A_h порядка n:

$$N \asymp 8n^3/3+4n^2+(C_1 l_1+C_2 l_2)n, \quad C_{1,2}=\mathrm{const}>0.$$

Первое слагаемое в этой оценке связано с разложением исходной матрицы на произведение ортогональных и двухдиагональной матриц при нахождении почти сингулярного разложения [216]. Второе слагаемое определяется окончательным вычислением приближения $z_{\eta\varkappa}$. Число l_1 — количество итераций, используемых при диагонализации матрицы в процессе нахождения сингулярного разложения [216]; l_2 — число итераций, используемых при определении $\lambda(h+\varkappa)$. Если $l_{1,2}\leqslant C_3 n$ ($C_3=\mathrm{const}>0$), что часто выполняется в практических расчетах, то $N\asymp 8n^3/3$. Такой же порядок числа операций имеют другие известные устойчивые методы решения СЛАУ (см., например, § 1 и [44, 55, 197]).

Представим результаты некоторых модельных расчетов по методу м.п.м. На рис. 24 приведены данные (точки) о точности восстановления $\Delta(h)$ с помощью принципа м.п.м. нормального псевдорешения модельной СЛАУ в зависимости от уровня h возмущения матрицы. Возмущения искусственно вводилось с помощью датчика равномерно распределенных псевдослучайных чисел. Точная матрица СЛАУ $[\bar{a}_{ij}]_{m\times n}$ имела вид $\bar{a}_{ij}=1$ ($i=1, \ldots, 16$, $j=1, \ldots, 15$), а точная правая часть $[\bar{u}_i]$ бралась в виде $u_1=16$, $u_i=15$ ($i=2, \ldots, 15$), $u_{16}=14$ и не возмущалась. Точное нормальное псевдорешение $[\bar{z}_j]$ задачи имеет вид $\bar{z}_j=1$. Оценка уровня возмущения h в данном случае вычислялась как $h=(mn)^{1/2}\cdot\max\{|a^h_{ij}-\bar{a}_{ij}|: i=1, \ldots, 16, j=1, \ldots, 15\}=\sqrt{240}\max\{|a^h_{ij}-\bar{a}_{ij}|\}$, где a^h_{ij} — элементы возмущенной матрицы. Величина h дана в процентах к $\|[\bar{a}_{ij}]\|=\sqrt{240}$. Точность восстановления $\Delta(h)=\max\{|z^h_j-\bar{z}_j|: j=1, \ldots, 15\}$ вычислялась по полученному с помощью программы при рассматриваемом h приближенному решению $[z^h_j]=\hat{A}^+_h\bar{u}$. При этом величина $\varkappa$ из (1) оказывалась пренебрежимо малой по сравнению с рассмотренными h. Величина $\Delta(h)$ на рис. 24 приведена в процентах к $\max\{|\bar{z}_j|: j=1, \ldots, 15\}=1$.

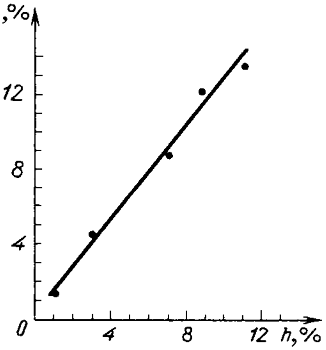

Рис. 24. Точность восстановления нормального псевдорешения модельной СЛАУ по методу м.п.м.

На рис. 25 показана зависимость времени t вычисления приближенного решения по принципу м.п.м. от размерности n задачи при фиксированном h. Зависимость получена при решении модельной задачи $[\bar{a}_{ij}]=[1]$, $[\bar{u}_i]=n$ $(i,\ j=1,\ 2, ..., n)$ с уровнем возмущения $h=1\%$, вычисляемым, как в предыдущей задаче. Зависимость включает в себя время расчета самой минимальной псевдообратной матрицы. Расчеты проводились на ЭВМ БЭСМ-6 (язык фортран).

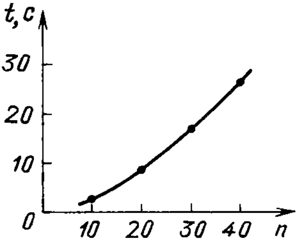

Рис. 25. Зависимость времени расчета приближенного решения по методу м.п.м. от размерности задачи

На рис. 26 показаны результаты решения сеточного аналога известной модельной задачи [206] с матрицей $[a_{ij}]$ $(i=1,\ 2, ..., m,$ $j=1,\ 2, ..., n)$: $a_{ij}=[1+100(x_i-s_j)^2]^{-1}$, $x_i=-2+4(i-1)/(m-1)$, $s_j=(j-1)/(n-1)$; $m=n=40$. Модельное точное решение $\bar{z}$ задавалось равенством, приведенным в [206]. Правая часть вычислялась по матрице и решению $\bar{z}$ в соответствии с (1.1). Матрица возмущалась случайными числами, равномерно распределенными на отрезке $[-0{,}01,\ 0{,}01]$. Этому соответствует оценка $h\leqslant 0{,}16$, которая использовалась в принципе м.п.м. Возмущения правой части имели уровень $\sigma=0{,}013$ (ср. [206]). Сплошная кривая — точное решение.

В заключение отметим, что метод м.п.м. весьма удобен при решении большого числа СЛАУ типа (1.1) с одной и той же

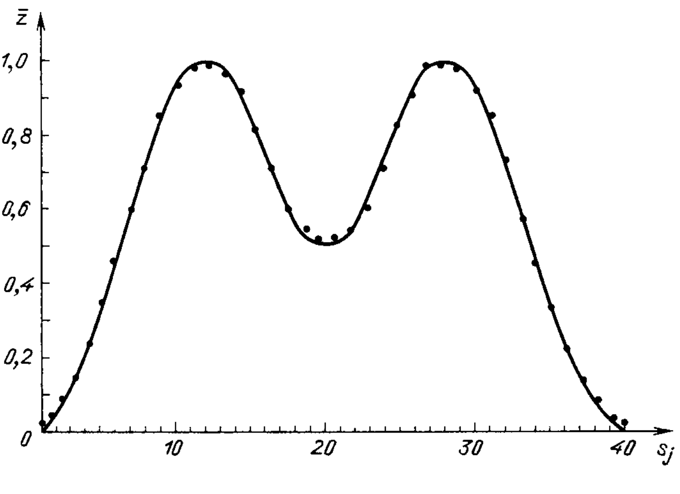

Рис. 26. Решение модельной задачи по методу м.п.м.: *1* — точное решение; *2* — приближенное решение

приближенной матрицей A_h и с различными правыми частями u_σ. Такая ситуация часто встречается при обработке больших серий экспериментальных данных.

ГЛАВА 6

ПРАКТИЧЕСКОЕ ПРИЛОЖЕНИЕ АЛГОРИТМОВ РЕШЕНИЯ НЕЛИНЕЙНЫХ НЕКОРРЕКТНЫХ ЗАДАЧ

В данной главе излагается опыт практического использования алгоритмов, теория которых дана в гл. 1—3, при решении конкретных прикладных задач. Большое число практических задач, для решения которых были использованы указанные регуляризующие алгоритмы, приводит к необходимости отбора лишь некоторых из таких задач. С другими примерами можно познакомиться в монографиях [19, 20, 71, 72, 202], а также в [92, 107, 138] и др. В главе рассматриваются также некоторые полезные для приложений модификации обобщенного метода невязки и обобщенного принципа невязки.

§ 1. Обратные задачи колебательной спектроскопии

Рассмотрим некоторые обратные задачи колебательной спектроскопии, заключающиеся в определении параметров силового поля молекулы по экспериментально измеряемым данным (в основном получаемым из анализа колебательных спектров молекул) [94].

Эти задачи приводят к операторным уравнениям вида

$$Az=u, \tag{1}$$

где $z\in Z$ — совокупность свойств объекта, подлежащих определению, $u\in U$ — совокупность экспериментальных данных, Z, U — некоторые пространства, A — оператор, определяемый выбором математической модели объекта.

В соответствии с [102] рассмотрим регуляризующие алгоритмы решения этих задач.

1. Основные уравнения. Постановка задачи. Понятие о силовом поле молекулы возникает из рассмотрения ее как квантово-механической системы, состоящей из точечных ядер и электронов, причем массы ядер и электронов существенно различны. В связи с этим может быть введен малый параметр ξ, равный отношению массы электрона к сумме масс ядер молекулы, и применена адиабатическая теория возмущения, основанная на разложении всех членов уравнения Шрёдингера для молекулы по степеням $\xi^{1/4}$ (строгий вывод может быть найден в [22]). Уравнения второго порядка теории возмущений дают возможность говорить о движении ядер в некотором эффективном силовом поле $\mathscr{U}(q_1, \ldots, q_n)$, создаваемом электронной подсистемой молекулы (здесь $q_1, \ldots, q_n$ — обобщенные координаты ядер).

Силовое поле играет важную роль в определении свойств молекулы. В частности, равновесная геометрическая конфигурация ядер атомов молекулы (если таковая существует) $q^0 = \{q_1^0, \ldots, q_n^0\}$ удовлетворяет соотношениям

$$(\partial \mathscr{U} / \partial q_i)|_{q^0} = 0, \quad i = 1, \ldots, n,$$

а свойства, связанные с малыми колебаниями молекулы, определяются матрицей силовых постоянных F:

$$F_{ij} = (\partial^2 \mathscr{U} / \partial q_i \partial q_j)|_{q^0}, \quad i, j = 1, \ldots, n.$$

Поскольку в дальнейшем будут использоваться экспериментальные данные по колебательному движению молекулы, то сформулируем задачу определения матрицы F, считая заданной равновесную конфигурацию q^0 (вопросы экспериментального определения последней рассмотрены, например, в [89]).

Число обобщенных координат, характеризующих конфигурацию из N ядер, должно быть $n \geqslant 3N-6$ ($3N-5$ в случае линейной молекулы); если $n > 3N-6$, то среди координат $q_1, \ldots, q_n$ окажутся зависимые. Их введение оправдывается удобствами учета симметрии молекул и интерпретации результатов.

Основным видом экспериментальной информации о колебаниях молекул являются частоты колебательных спектров. Они связаны с матрицей силовых постоянных F уравнением на собственные значения

$$GFL = L\Lambda, \tag{2}$$

где Λ — диагональная матрица, состоящая из квадратов частот нормальных колебаний молекулы, $\Lambda = \mathrm{diag}(\omega_1^2, \ldots, \omega_n^2)$, G — матрица кинетической энергии в импульсном представлении, зависящая только от масс ядер и их равновесной конфигурации, которую будем считать известной (возможно, с некоторой погрешностью), L — матрица собственных векторов.

В рамках рассматриваемого приближения силовое поле молекулы не зависит от масс ядер, поэтому при наличии спектров m изотопных разновидностей молекулы вместо (2) используется система

$$G_i F L_i = L_i \Lambda, \tag{3}$$

где индекс i нумерует изотопные разновидности.

Важную информацию о силовом поле молекулы представляют постоянные Кориолиса, характеризующие колебательно-вращательное взаимодействие в молекуле. Они определяются по расщеплению вырожденных колебательных уровней и связаны с матрицей F через собственные векторы задачи (2):

$$\zeta = \frac{1}{M^2} L^* \mathscr{A} \mathfrak{M} \mathscr{A}^* L, \tag{4}$$

где ζ — матрица с векторными элементами, состоящая из кориолисовых постоянных, $\mathfrak{M}$ — диагональная матрица, состоящая из масс ядер, M — сумма масс ядер молекулы, $\mathscr{A}$ — матрица, связывающая декартовы смещения атомов Δr с введенными координатами q:

$$\Delta r_i = \sum_{k=1}^{n} \mathscr{A}_{ik} q_k, \quad i = 1, \ldots, N,$$

а произведение векторных элементов в (4) понимается в смысле векторного произведения. Матрица $\mathscr{A}$ может быть определена по равновесной конфигурации молекулы.

Методы газовой электронографии позволяют определить среднеквадратичные амплитуды колебаний межъядерных расстояний, квадраты которых являются диагональными элементами матрицы

$$\mathscr{D} = HL\Delta L^* H^*. \tag{5}$$

Здесь Δ — диагональная матрица с элементами

$$\Delta_i = \frac{\hbar}{2\omega_i} \operatorname{cth} \frac{\hbar\omega_i}{kT}, \quad i = 1, \ldots, n,$$

где $\hbar$ — постоянная Планка, k — постоянная Больцмана, T — температура, ω_i есть i-я частота колебательного спектра; H — матрица с элементами

$$H_{mk} = |R_i^0 - R_j^0|^{-1} (R_i^0 - R_j^0, \ \mathscr{A}_{ik} - \mathscr{A}_{jk}),$$

где i, j — индексы атомов, соответствующих m-му межатомному расстоянию, R_i^0 — радиус-вектор i-го ядра в равновесной конфигурации. Матрица H зависит лишь от геометрии молекулы и от масс атомов.

Точно так же могут быть выписаны зависимости других измеряемых величин (в частности, постоянных центробежного искажения) от матрицы F (через собственные частоты ω и матрицу L, являющиеся функциями элементов матрицы F).

Совокупность уравнений (2)—(5) (или части их в зависимости от имеющихся экспериментальных данных) будем рассматривать как одно операторное уравнение с оператором A, ставящим в соответствие действительной симметричной матрице F набор из собственных значений задачи (2), кориолисовых постоянных ζ, амплитуд $\mathscr{D}$ и т. д.

Этот набор данных представим вектором из конечномерного пространства U. Введем также вектор z из конечномерного пространства Z, состоящий либо из элементов матрицы F, либо из величин, которыми эта матрица параметризована. Приходим к формулировке обратной задачи в терминах гл. 2.

Задача I. Дано уравнение (1), где $z \in D \subseteq Z$, $u \in U$; Z, U — конечномерные пространства, D — замкнутое множество априорных ограничений задачи, A — нелинейный, непрерывный на D оператор. Требуется найти приближенное решение уравнения (1), если вместо A, u заданы их приближения A_h, u_σ такие, что $\|u - u_\sigma\| \leqslant \sigma$, $\|Az - A_h z\| \leqslant \varphi[h, z]$ для всех $z \in D$. Здесь $\varphi[h, z]$ — известный непрерывный функционал, стремящийся к нулю при $h \to 0$ равномерно для всех $z \in D \bigcap \bar{S}(0, R)$, где $\bar{S}(0, R)$ — замкнутый шар с центром в $z=0$ и с некоторым радиусом R.

Погрешность задания оператора A связана с погрешностью определения равновесной конфигурации молекулы, параметры которой находятся из эксперимента. Вид функционала φ будет уточнен ниже.

Заметим, что задача I может не удовлетворять ни одному из следующих условий корректности задачи.

а. *Разрешимость.* Легко видеть, что, например, система (3) совместна при условии

$$\det G_i / \det \Lambda_i = \text{const}, \quad i=1, \ldots, m,$$

которое может нарушаться как в силу неточного измерения частот Λ_i, так и за счет неточного задания геометрии молекулы G_i, а при точно заданных данных — за счет ангармоничности колебаний, не учитываемой оператором задачи I.

б. *Единственность решения задачи.* Если имеется лишь один спектр, то задача I сводится к обратной задаче на собственные значения (2), откуда при невырожденности G следует, что решением задачи I является всякая матрица F такая, что

$$F = G^{-1/2} C \Lambda C^* G^{-1/2}, \tag{6}$$

где C — произвольная ортогональная матрица.

в. *Устойчивость решения относительно возмущений u, A.* Пример неустойчивости легко строится для системы вида (3). Этот факт хорошо известен спектроскопистам.

Возникает вопрос, что считать решением задачи I. Рассмотрим задачу об отыскании нормального псевдорешения задачи I.

З а д а ч а II. Найти

$$\operatorname{Arg}\inf\{\|z-z_0\|: z \in D, \ \|Az-u\| = \mu\},$$

где $\mu = \inf\{\|Az-u\|: \ z \in D\}$.

Элемент $z_0 \in D$ должен быть задан из априорных соображений о решении, использующих как приближенные квантово-механические расчеты, так и другие идеи (например, переносимость силовых постоянных в сходных фрагментах молекул).

Очевидно, в случае существования и единственности решения задачи I его нормальное псевдорешение совпадает с решением.

С учетом всего изложенного поставим следующую задачу.

З а д а ч а III. Пусть дано уравнение (1). Требуется по приближенным данным $\{A_h, u_\sigma, h, \sigma\}$ найти такие приближения $z_\delta \in D$ к решению $\bar{z}$ задачи II, что

$$\lim_{\delta \to 0} \|z_\delta - \bar{z}\| = 0, \quad \delta \equiv (h, \ \sigma),$$

т. е. алгоритм отыскания z_δ должен быть регуляризующим.

2. Обратная колебательная задача. Рассмотрим простейшую постановку задачи I (задача I′). Известен колебательный спектр одной молекулы, и в уравнении (1) оператор A ставит в соответствие вектору $z \in \mathbf{R}^{n(n+1)/2}$, составленному из элементов симметричной матрицы F порядка n, упорядоченный набор собственных значений матрицы GF. В качестве правой части $u \in \mathbf{R}^n$ будем использовать упорядоченный набор квадратов частот колебаний молекулы.

Задача II′. Найти нормальное решение

$$\operatorname{Arg} \inf \{ \|z - z_0\| : z \in Z, \ Az = u \},$$

при условии, что задача I′ всегда разрешима, и притом неоднозначно (кроме случая $n = 1$).

Поскольку оператор A в (1) полностью определяется заданием матрицы кинематических коэффициентов G, будем оценивать уклонение приближенно заданного оператора A_h (соответствующего некоторой G_ξ) от точного A (соответствующего G) погрешностью задания матрицы G. Пусть в некоторой матричной норме $\|G - G_\xi\| \leqslant \xi$. В пространстве $\mathbf{R}^n$ правых частей u введем евклидову норму с весами. Пусть вместо точного значения правой части u задана u_σ такая, что $\|u - u_\sigma\| \leqslant \sigma$.

Справедливы следующие теоремы об устойчивости задач I′, II′ (см. [102]).

Теорема 1. *Задача* I′ *устойчива в псевдометрике Хаусдорфа относительно возмущений оператора и правой части.*

Доказательство. Пусть Z_0 — множество решений задачи I′ при точно заданных, а Z_η — при приближенно заданных операторе A и правой части u. Требуется доказать, что

$$\lim_{\eta \to 0} P(Z_\eta, Z_0) = 0, \quad \eta \equiv (\xi, \sigma),$$

где $P(Z_\eta, Z_0)$ — расстояние между множествами Z_η и Z_0, равное

$$P(Z_\eta, Z_0) = \sup_{x \in Z_\eta} \inf_{z \in Z_0} \|x - z\| + \sup_{z \in Z_0} \inf_{x \in Z_\eta} \|x - z\|.$$

Воспользуемся известным представлением решения задачи I′, а именно формулой (6), где $\Lambda = \operatorname{diag}(u_1, \ldots, u_n)$, C — произвольная ортогональная матрица. Аналогично $F_\eta = G_\xi^{-1/2} K \Lambda_\sigma K^* G_\xi^{-1/2}$, где K также ортогональна. Тогда для некоторой матричной нормы

$$\begin{aligned}
\inf_{F_\eta} \|F - F_\eta\| &= \inf_{K} \|G^{-1/2} C \Lambda C^* G^{-1/2} - G_\xi^{-1/2} K \Lambda_\sigma K^* G_\xi^{-1/2}\| \leqslant \\
&\leqslant \|G^{-1/2} C \Lambda C^* G^{-1/2} - G_\xi^{-1/2} C \Lambda_\sigma C^* G_\xi^{-1/2}\| \leqslant \\
&\leqslant \|G_\xi^{-1/2}\| \cdot \|G^{-1/2}\| \cdot \|\Lambda - \Lambda_\sigma\| + \|G^{-1/2} - \\
&\qquad - G_\xi^{-1/2}\| \cdot (\|\Lambda\| \cdot \|G^{-1/2}\| + \|\Lambda_\sigma\| \cdot \|G_\xi^{-1/2}\|).
\end{aligned}$$

Учитывая, что $\|\Lambda_\sigma\| \leqslant \|\Lambda\| + \sigma$ и что для всех ξ из некоторого отрезка $[0, \xi_0]$ справедливы оценки

$$\|G_\xi^{-1}\| \leqslant \|G^{-1}\| / (1 - \xi \|G^{-1}\|), \quad \|G^{-1/2} - G_\xi^{-1/2}\| \leqslant f(\xi),$$

где $f(\xi)$ — непрерывная монотонно возрастающая функция, $f(0)=0$, получаем

$$\inf_{F_\eta} \| F-F_\eta \| \leqslant f(\xi) \| \Lambda \| \cdot \| G^{-1/2} \| +$$
$$+\left[\sigma \| G^{-1/2} \| + f(\xi)(\| \Lambda \| + \sigma)\right] \cdot \| G^{-1} \|^{1/2} / (1-\xi \| G^{-1} \|)^{1/2}.$$

Пользуясь эквивалентностью норм для z и F, а также u и Λ, имеем

$$\sup_{z\in Z_0} \inf_{x\in Z_\eta} \| x-z \| \leqslant k_1 f(\xi)+k_2\sigma+k_3 f(\xi)\sigma,$$

где k_1, k_2, k_3 — некоторые константы. Проделав то же для второго слагаемого в выражении для метрики Хаусдорфа, окончательно получим

$$P(Z_\eta,\ Z_0) \leqslant k'_1 f(\xi)+k'_2\sigma+k'_3 f(\xi)\sigma \to 0,$$
$$\xi,\ \sigma \to 0.$$

Теорема 2. *Задача* II′ *устойчива относительно возмущений правой части и оператора.*

Здесь в случае неединственности нормального решения задачи I′ имеется в виду β-сходимость приближенных решений ко множеству нормальных решений точной задачи.

Доказательство основано на том, что при любом η в η-окрестности $\bar z$ найдется элемент $\tilde z \in Z_\eta$, который в свою очередь обладает свойством нормальности по отношению к z_0 (если бы последнее было несправедливо, то в окрестности $\bar z_\eta$ — решения задачи с приближенными данными — нашелся бы элемент из Z_0 с меньшим, чем у $\bar z$, расстоянием до z_0). Существенно, что Z_0, Z_η — замкнутые, ограниченные в совокупности множества.

3. Регуляризующие алгоритмы для решения задачи отыскания нормального псевдорешения. Рассмотрим задачу I в гильбертовом пространстве (задача I″). Требуется по заданному набору данных A_h, u_σ, h, σ построить приближения z_δ к решению $\bar z$ этой задачи такие, что $z_\delta \to \bar z$ при $\delta \equiv (h,\ \sigma) \to 0$. Для решения этой задачи могут быть использованы методы, описанные в гл. 1—3.

Ниже предлагается модификация обобщенного принципа невязки, основанная на возможности оценки погрешности оператора в виде

$$\| Az-A_h z \| \leqslant h \| A_h z \| \equiv \varphi\left[h,\ z\right], \tag{7}$$

что соответствует заданию относительной погрешности Az и является для рассматриваемой задачи более удобной оценкой, чем монотонная функция вида $\psi(h,\ \| z-z_0 \|)$.

Пусть z^α — экстремаль (возможно, неединственная) сглаживающего функционала

$$M^\alpha\left[z\right] = \| A_h z-u_\sigma \|^2+\alpha \| z-z_0 \|^2 \tag{8}$$

на множестве D. Существование экстремали следует из результатов гл. 2.

Введем функцию

$$\rho(\alpha)=\| A_h z^{\alpha}-u_\sigma \| - \frac{1}{1-h}\left[\hat{\mu}_\delta+k(\sigma+h\| u_\sigma \|)\right], \tag{9}$$

где $k>1$ — константа;

$$\hat{\mu}_\delta=\inf\{\| A_h z-u_\sigma \|+\sigma+h\| A_h z\| : z\in D\}.$$

Если выполнено условие

$$\| A_h z_0-u_\sigma \| > \frac{1}{1-h}\left[\hat{\mu}_\delta+k(\sigma+h\| u_\sigma \|)\right], \tag{10}$$

то уравнение $\rho(\alpha)=0$ имеет обобщенное решение $\alpha_\delta>0$ (т. е. α_δ таково, что $\rho(\alpha)>0$ при $\alpha>\alpha_\delta$, $\rho(\alpha)<0$ при $\alpha<\alpha_\delta$). Если α_δ — точка непрерывности $\rho(\alpha)$, то $\rho(\alpha_\delta)=0$. Это утверждение следует из монотонности $\rho(\alpha)$ и из предельных соотношений (при $\alpha\to+0$, $\alpha\to+\infty$).

Сформулируем алгоритм отыскания приближений к нормальному решению задачи I″.

Шаг 1. Если условие (10) не выполнено, полагаем $z_\delta=z_0$.

Шаг 2. Если условие (10) выполнено, находим $\alpha_\delta>0$ — обобщенное решение уравнения $\rho(\alpha)=0$ — и полагаем $z_\delta=z^{\alpha_\delta}$. Если экстремаль функционала (8) неединственна, выбираем такую из них, что

$$\| A_h z^{\alpha_\delta}-u_\sigma \| \leqslant \frac{1}{1-h}\left[\hat{\mu}_\delta+k(\sigma+h\| u_\sigma \|)\right]$$

(возможность такого выбора обосновывается, как в § 2.2).

Докажем лемму, аналогичную теореме 1.3.1.

Лемма 1. *Пусть выполнены условия задачи* I″ *и*

$$\hat{\mu}_\delta\equiv\hat{\mu}(A_h,\ u_\sigma)=\inf\{\| A_h z-u_\sigma \|+\sigma+\varphi[h,\ z] : z\in D\}.$$

Тогда $\hat{\mu}(A_h,\ u_\sigma)\geqslant\mu$, *причем*

$$\lim_{\delta\to 0}\hat{\mu}_\delta=\mu. \tag{11}$$

Доказательство. Поскольку для любого $z\in D$

$$\mu=\inf\{\| Az-u\| : z\in D\}\leqslant\| Az-u\|\leqslant\| A_h z-u_\sigma \|+\sigma+\varphi[h,\ z],$$

то

$$\mu\leqslant\inf\{\| A_h z-u_\sigma \|+\sigma+\varphi[h,\ z] : z\in D\}\equiv\hat{\mu}_\delta.$$

Далее,

$$\hat{\mu}_\delta\leqslant\| A_h\bar{z}-u_\sigma \|+\sigma+\varphi[h,\ \bar{z}]\leqslant\mu+2(\sigma+\varphi[h,\ \bar{z}]),$$

поэтому

$$\mu\leqslant\hat{\mu}_\delta\leqslant\mu+2(\sigma+\varphi[h,\ \bar{z}]),$$

откуда получается утверждение (11).

Теорема 3 [102]. *Сформулированный алгоритм является регуляризующим.*

Доказательство. Пусть $\bar{z}=z_0$, тогда легко показать, что условие (10) не выполнено и $z_\delta=z_0$. Если же $\bar{z}\neq z_0$, то $\|Az_0-u\|=\mu+\varepsilon$ $(\varepsilon>0)$. Отсюда

$$\|A_h z_0-u_\sigma\|\geqslant\frac{\mu+\varepsilon-(\sigma+h\|u_\sigma\|)}{1+h}. \tag{12}$$

Правая часть (12) при h, $\sigma\to 0$ стремится к $\mu+\varepsilon$, а правая часть (10) — к μ; поэтому, начиная с некоторых h, σ, условие (10) будет выполнено.

Рассмотрим $\alpha_\delta>0$ — обобщенное решение уравнения $\rho(\alpha)=0$. Пусть экстремаль z^{α_δ} функционала (8) неединственная. Рассмотрим $z_+^{\alpha_\delta}$ — такую из них, что

$$\|A_h z_+^{\alpha_\delta}-u_\sigma\|\geqslant\frac{1}{1-h}\left[\hat{\mu}_\delta+k(\sigma+h\|u_\delta\|)\right].$$

Из условия $M^{\alpha_\delta}[z^{\alpha_\delta}]\leqslant M^{\alpha_\delta}[\bar{z}]$ имеем

$$\xi(h,\ \sigma)\equiv\frac{[\hat{\mu}_\delta+k(\sigma+h\|u_\sigma\|)]^2-(\hat{\mu}_\delta+h\|u_\sigma\|+\sigma)^2}{(1-h)^2\alpha_\delta}\leqslant\|\bar{z}-z_0\|^2-\|z_+^{\alpha_\delta}-z_0\|^2,$$

а также $\|z_+^{\alpha_\delta}-z_0\|\leqslant\|\bar{z}-z_0\|$. Можно показать, что

$$\lim_{h,\ \sigma\to 0}\xi(h,\ \sigma)=0.$$

Рассматривая экстремаль $z_-^{\alpha_\delta}$ такую, что

$$\|A_h z_-^{\alpha_\delta}-u_\sigma\|\leqslant\frac{1}{1-h}\left[\hat{\mu}_\delta+k(\sigma+h\|u_\sigma\|)\right],$$

и пользуясь соотношениями $\xi(h,\ \sigma)\to 0$ при h, $\sigma\to 0$ и

$$M^{\alpha_\delta}[z_-^{\alpha_\delta}]\leqslant M^{\alpha_\delta}[\bar{z}],$$

получаем

$$\overline{\lim_{\delta\to 0}}\|z_-^{\alpha_\delta}-z_0\|\leqslant\|\bar{z}-z_0\|.$$

Выделяя из $\{z_-^{\alpha_\delta}\}$ слабо сходящуюся подпоследовательность и учитывая, что

$$\lim_{\delta\to 0}\|Az_-^{\alpha_\delta}-u\|=\mu,$$

получаем слабую сходимость z^{α_δ} к $\bar{z}$. При использовании схемы компактного вложения (или требуя усиленной непрерывности A, A_h) можно получить результат теоремы.

Для вариантов задачи I″, когда оценка погрешности оператора A не может быть записана в виде (7), но выполнены требования задачи I″, можно использовать следующий вариант обобщенного метода невязки (рассмотрим случай конечномерных Z, U).

Задача IV. Найти

$$\operatorname{Arg}\inf\{\|z-z_0\|:\ z\in Z_\delta\},$$
$$Z_\delta\equiv\{z\in D:\ \|A_h z-u_\sigma\|\leqslant\sigma+\varphi[h,\ z]+\hat{\mu}_\delta\},$$

где оценка меры несовместности точной задачи сверху есть

$$\hat{\mu}_\delta=\inf\{\|A_h z-u_\sigma\|+\sigma+\varphi[h,\ z]:\ z\in D\}.$$

Лемма 2. *Задача* IV *разрешима для любого* $u_\sigma\in U$ *такого, что* $\|u-u_\sigma\|\leqslant\sigma$, *и для непрерывного оператора* A_h *такого, что* $\|A_h z-Az\|\leqslant\varphi[h,\ z]$ *при всех* $z\in D$.

Доказательство леммы следует из непустоты Z_δ при всех δ (так как $\bar{z}\in Z_\delta$), а также из его ограниченности и замкнутости. Ограниченность множества, на котором производится минимизация $\|z-z_0\|$, достигается путем его замены на $Z_\delta\cap\bar{S}(z_0,\ \|\bar{z}-z_0\|)$.

Теорема 4. *Алгоритм, определяемый экстремальной задачей* IV, *является регуляризующим для задачи* I″.

Доказательство следует из того, что $\|z_\delta-z_0\|\leqslant\|\bar{z}-z_0\|$ для всех δ, и из непрерывности $\|Az_\delta-u\|$. В бесконечномерных пространствах следует потребовать выпуклости множества D, а также усиленной непрерывности A, A_h (либо использовать схему компактного вложения).

Рассмотренные алгоритмы в случае неединственности решения задачи I″ дают β-сходимость полученных приближений ко множеству решений I″.

4. Оценка погрешности оператора. Для оператора обратной колебательной задачи (соответствующего уравнениям (2), (3)) возможно получение оценки погрешности в виде (7). Одновременно имеет место оценка

$$\|Az-A_h z\|\leqslant h_1\|z\|, \tag{13}$$

где h, $h_1\to 0$ при уменьшении погрешностей равновесной конфигурации молекулы.

Пусть вместо точной матрицы G (см. п. 1) известна G_ξ, такая, что $\|G-G_\xi\|\leqslant\xi$ в некоторой матричной норме. Пусть матрица G соответствует точному оператору A, а G_ξ — приближенно заданному A_h.

Согласно известной теореме [113] о возмущении собственных значений простой матрицы A, имеем

$$|\tilde{\mu}_j-\mu_j|\leqslant\|B\|\max\{\nu(A),\ \nu(C)\},$$

если $C=A+B$ — простая матрица, μ_j — собственные значения матрицы A, $\tilde{\mu}_j$ — собственные значения C, $\nu(A)$ — известная скалярная функция матрицы A. Применяя теорему в случае $A=GF$, $C=G_\xi F$, получаем

$$|\tilde{\mu}_j-\mu_j|\leqslant\|F\|\cdot\|G-G_\xi\|\cdot\max\{\varkappa^{1/2}(G),\ \varkappa^{1/2}(G_\xi)\},$$

где $\nu(G)$ — число обусловленности матрицы G. Учитывая, что

$$\|Az-A_h z\|^2=\sum_j|\tilde{\mu}_j-\mu_j|^2\cdot p_j,$$

где p_j — некоторые положительные веса, получаем утверждение (13), а (7) будет следовать из него в силу формулы (6).

5. Численная реализация методов. Ввиду устойчивости задачи II′ (см. п. 2) для ее решения могут быть применены просто реализуемые на ЭВМ методы типа метода Монте-Карло. Такой подход описан в [101]; здесь мы на нем останавливаться не будем.

Для решения задач, описанных в п. 3, применялись численные методы. Решение задачи, связанное с минимизацией сглаживающего функционала (8), производилось с помощью метода сопряженных градиентов, модифицированного на случай неквадратичного функционала, при учете априорных ограничений на решение вида $z_i^* \leqslant z_i \leqslant z_i^{**}$, где z_i^*, z_i^{**} — некоторые константы.

Один шаг минимизации функционала $f(z)$ имеет следующий вид. Пусть вычислено $z^{(k)}$, находящееся внутри области ограничений. Тогда

$$s^{(k)} = -f'\left[z^{(k)}\right] + s^{(k-1)} \| f'\left[z^{(k)}\right] \|^2 / \| f'\left[z^{(k-1)}\right] \|^2,$$

$$\lambda_{\max} = \min_{\substack{i \in I_+ \\ j \in I}} (\lambda_i, \ \mu_j), \quad I_+ = \{ i: \ s_i^{(k)} > 0 \},$$

$$I_- = \{ i: \ s_i^{(k)} < 0 \}, \ \lambda_i = \frac{-z_i^* + z_i^{(k)}}{s_i^{(k)}}, \quad \mu_j = \frac{z_j^{(k)} - z_j^{**}}{s_j^{(k)}}.$$

Далее решается задача

$$\operatorname{Arg} \min \{ f\left(z^{(k)} + \lambda s^{(k)}\right): \lambda \in [0, \ \lambda_{\max}] \}.$$

Если $\lambda^* > 0$ — ее решение, то $z^{(k+1)} = z^{(k)} + \lambda^* s^{(k)}$.

Если $z^{(k+1)}$ принадлежит границе области ограничений, то на следующем шаге минимизация по тем компонентам z, для которых $z_i^{(k+1)} = z_i^*$ или $z_i^{(k+1)} = z_i^{**}$, не производится.

В результате размерность множества, на котором производится минимизация, сокращается, пока не будет найден минимум на каком-либо подпространстве Z. В дальнейшем в зависимости от знака $f'_i(z)$ компонента z_i может быть вновь включена в процесс минимизации.

Для обеспечения убывания функционала $f(z)$ по направлению s периодически полагается $s = -f'(z)$. Это происходит также при любом изменении размерности множества, на котором производится минимизация.

Решение уравнения $\rho(\alpha) = 0$ (см. (9)) производится путем уменьшения α от некоторого значения $\alpha_0 > 0$, для которого $\rho(\alpha_0) > 0$.

Для решения экстремальной задачи с невыпуклыми ограничениями IV использовался метод линеаризации [172], в данном случае применяемый к задаче

$$\operatorname{Arg} \inf \{ \| z - z_0 \|^2: z \in D, \ f(z) \leqslant 0 \},$$

где $f(z)$ определяется постановкой задачи IV.

Один шаг алгоритма имеет следующий вид: пусть известно $z^{(k)} \in D$; тогда решаются задачи

$$\operatorname{Arg} \inf_p \{ 2\left(z^{(k)} - z_0, \ p\right) + \frac{1}{2}(p, \ p): \left(f'\left(z^{(k)}\right), \ p\right) + f\left(z^{(k)}\right) \leqslant 0 \}, \tag{14}$$

$$\operatorname{Arg} \inf_\lambda \{ \| z^{(k)} - z_0 + \lambda p_k \|^2 + N \mathscr{F}\left(z^{(k)} + \lambda p_k\right): \ \lambda \geqslant 0 \}, \tag{15}$$

где p_k — решение задачи (14), $\mathscr{F}(z) = \max\{0; f(z)\}$; если λ^* — решение (15), то полагается $z^{(k+1)} = z^{(k)} + \lambda^* p_k$.

При этом задача (14) решается аналитически, а именно если $\bar{p} \equiv -2(z^{(k)} - z_0)$ удовлетворяет ограничениям, то $p_k = \bar{p}$, иначе

$$p_k = \bar{p} - \frac{f(z^{(k)}) + (f'(z^{(k)}), \bar{p})}{\| f'(z^{(k)}) \|^2} f'(z^{(k)}).$$

Коэффициент при $-f'(z^{(k)})$ является множителем Лагранжа для задачи; его явное вычисление позволяет выбирать параметр N алгоритма.

На основе описанных алгоритмов создан комплекс программ обработки спектроскопических данных на ЭВМ. Он включает блок чтения информации, блок подготовки вспомогательных данных, блок решения обратной задачи, блок решения прямой задачи, программу-диспетчер. Подробно работа комплекса программ описана в [97]. Отметим кратко основные этапы решения задачи.

Вначале должна быть введена подготовленная пользователем информация о равновесной геометрической конфигурации молекулы и о ее симметрии, а также об имеющихся экспериментальных данных. Затем рассчитываются вспомогательные матрицы G, $\mathscr{A}$, H (см. п. 1). Далее производится решение задачи I″ с помощью минимизации функционала (8) для разных значений параметра α и выбор этого параметра или же решение задачи IV с помощью метода линеаризации. Интересующие пользователя результаты выводятся на печать.

В качестве примера приведем некоторые данные о расчете матрицы силовых постоянных для молекулы CH_3SiH_3 и ее изотопных модификаций CH_3SiD_3, CD_3SiH_3, CD_3SiD_3. Из 50 независимых параметров силового поля 16 полагались равными нулю из физических соображений; экспериментальные данные состояли из 48 частот колебаний (по 12 на каждую изотопную модификацию). В соответствии со сказанным в п. 1 задача оказывается несовместной.

В результате отыскания приближения к нормальному псевдорешению задачи получена матрица силовых постоянных, воспроизводящая экспериментальную информацию с точностью 1—5% и ближайшая к априорно заданной диагональной матрице F_0 (максимальный недиагональный элемент не превосходил 1/4, а в большинстве случаев 1/10 величины диагональных элементов). Более подробно о расчетах с применением данного комплекса программ говорится в [100, 235, 236].

§ 2. Оптимизация режима лекарственной терапии

Проблема дозировки лекарств является одной из важнейших в практической медицине. Особенно это относится к лекарственным препаратам с «узким коридором» действия, когда даже небольшая передозировка может привести к тяжелым побочным явлениям в организме, а малая недозировка ведет к отсутствию терапевтического эффекта. Дозировка лекарства существенно зависит от индивидуальных особенностей организма пациента (пол, возраст, масса тела, вид патологии и др.). Вместе с тем инструкции по применению лекарственных препаратов, особенно недавно разработанных, в ряде случаев не учитывают детально этих особенностей.

Многие препараты, предназначенные для лечения острых инфекций, являются в определенной мере токсичными. Поэтому необходимо, чтобы дозировка таких препаратов не только удовлетворяла условиям оптимального воздействия препарата, а также по возможности приводила к минимальным его затратам. К этим же соображениям приводит и тот факт, что широкое, недостаточно обоснованное применение больших доз нового препарата приводит часто к быстрому развитию устойчивости возбудителей заболевания и снижению эффективности препарата.

Сказанное приводит к постановке математической задачи оптимизации дозировки лекарственных препаратов. Один из ее возможных вариантов [95] дается ниже.

Пусть функция $c(t)$, зависящая от времени t, имеет смысл терапевтической концентрации препарата в очаге воспаления, а функция $u(t)$ представляет собой зависимость вводимых в организм доз лекарства в каждый момент времени. Принимая для организма больного простейшую модель «черного ящика», можно получить связь «входного воздействия» u с «выходной функцией» c:

$$c(t)=\int_0^t K(t-\tau)u(\tau)\,d\tau, \quad 0\leqslant t\leqslant T. \tag{1}$$

Функция $K(t)$ — весовая (передаточная) функция организма как «черного ящика», представляющая собой «ответ» организма на δ-образное воздействие. По смыслу задачи функции $K(t)$ и $c(t)$ непрерывны и неотрицательны при $t\geqslant 0$, причем $K(0)=0$.

Для рассматриваемой задачи характерным является класс функций $u(t)$ следующего вида:

$$u(t)=\sum_{k=1}^{N} u_k\delta(t-t_k), \tag{2}$$

где N — заданное количество дозировок, t_k — время введения лекарственного вещества, $0<t_1\leqslant\ldots\leqslant t_N\leqslant T_0<T$, $T_0=\text{const}>0$, $u_k\geqslant 0$ — величина вводимой в момент t_k дозы лекарства. Специфическая форма (2) входных функций $u(t)$ с медицинской точки зрения обусловлена тем, что время растворения таблетки в желудочно-кишечном тракте оценивается секундами; время же выведения препарата из организма после однократного введения составляет несколько часов.

Функция $K(t)$ может быть определена экспериментально как реакция данного специфического организма на однократный прием единичной дозы препарата. Тогда связь $c(t)$, $u(t)$ вида (2) принимает форму

$$c(t)=\sum_{k=1}^{N} K(t-t_k)u_k. \tag{3}$$

Во многих случаях, исходя из специфических медицинских требований, можно установить вид функции $c(t)$, соответствующий необходимому терапевтическому эффекту. При этом возникает возможность решения обратной задачи: по заданной $c(t)$ найти такие времена t_k и дозы u_k, чтобы выполнялось условие (3). Уравнение (3), как правило, не разрешимо относительно (t_k, u_k) $(k=1, ..., N)$ при указанных ограничениях на неизвестные. Поэтому вместо уравнения (3) следует рассмотреть экстремальную задачу: найти такие времена t_k^* $(0<t_1^*\leqslant...\leqslant t_N^*\leqslant T_0)$ и дозы $u_k^*\geqslant 0$, что

$$\Phi(t_1^*, ..., t_N^*;\ u_1^*, ..., u_N^*)=$$

$$=\inf\{\Phi(t_1, ..., t_N;\ u_1, ..., u_N):\ 0<t_1\leqslant...\leqslant t_N\leqslant T_0;\ u_k\geqslant 0\}, \tag{4}$$

где

$$\Phi(t_1, ..., t_N;\ u_1, ..., u_N)=\left\|\sum_{k=1}^{N} K(t-t_k)\,u_k-c(t)\right\|^2_{L_2[0,T]}.$$

Можно показать, что задача (4) разрешима при условии выполнения в некоторой точке $t^*\in(0, T-T_0)$ неравенства $K(t^*)>0$ (см. [95]).

Иногда вследствие требований больничного режима моменты введения t_k фиксированы. Тогда можно рассматривать другую экстремальную задачу: найти такие дозы $u_k^{**}\geqslant 0$, для которых

$$\Phi_0(u_1^{**}, ..., u_N^{**})=\inf\{\Phi_0(u_1, ..., u_N):\ u_k\geqslant 0,\ \ k=1, ..., N\}, \tag{5}$$

где $\Phi_0(u_1, ..., u_N)\equiv\Phi(t_1, ..., t_N;\ u_1, ..., u_N)$, $t_k=\mathrm{const}$. Задача (5) также разрешима при указанном выше условии на функцию $K(t)$.

Поскольку задачи (4), (5) могут быть неустойчивы по отношению к возмущениям данных (K, c), а также могут быть неоднозначно разрешимы, для их решения можно использовать принцип выделения Ω-оптимальных решений (см. гл. 1). Для этого можно ввести стабилизирующий функционал Ω:

$$\Omega(t_1, ..., t_N;\ u_1, ..., u_N)=\sum_{k=1}^{N}\left[p_k(t_k-t_k^0)^2+q_k(u_k-u_k^0)^2\right], \tag{6}$$

в котором числа t_k^0, u_k^0 характеризуют некоторый апробированный режим введения лекарственного препарата. Весовые множители $p_k\geqslant 0$, $q_k>0$ фиксированы. Для задачи (5) можно положить $p_k=0$. Функционал (6) определяет меру отклонения режима введения (2) с параметрами t_k, u_k от известного режима с параметрами t_k^0, u_k^0. Решая задачу о поиске Ω-оптимальных решений задач (4), (5), можно найти режимы введения лекарства, не только обеспечивающие наибольшую близость терапевтической концентрации, которая создается в очаге инфекции согласно модели (3), к заданной функции $\bar{c}(t)$, но и наименее уклоняющиеся от апробированных на практике режимов.

С другой стороны, если для задачи (5) принять

$$\Omega(u_1, ..., u_N)=\sum_{k=1}^{N} q_k(u_k)\,u_k, \tag{7}$$

$$q_k(u)\geqslant q_0=\mathrm{const}>0,$$

то в зависимости от выбора функции $q_k(u)$ Ω-оптимальные решения задачи (5) будут характеризовать режим введения с наименьшим токсическим воздействием на организм (в этом случае $q_k(u)$ берется из эксперимента) или же, например, с минимальной стоимостью затраченного препарата (если $q_k(u)=q_0=\text{const}$).

Методика оптимизации дозировки лекарственных препаратов, основанная на решении некорректных экстремальных задач (4), (5), была реализована на базе практического лечебного учреждения применительно к лечению больных с острыми воспалительными заболеваниями почек (пиелонефритом). Для этих целей использовалась ПЭВМ. Написанная для нее программа поиска Ω-оптимальных решений задач (4), (5) основывалась на алгоритме о.п.н. для экстремальных задач (см. гл. 1). В качестве базового метода минимизации использовался метод Флетчера — Ривса (см. [170]). Параметр регуляризации определялся по обобщенному принципу невязки в результате серии решений типичных задач и в дальнейшем фиксировался.

Основным условием успешного лечения воспалительного процесса в почках у больных является создание и поддержание постоянства концентрации препарата на заданном уровне. Например, считается, что при лечении препаратом 5-НОК наилучшее подавление возбудителей пиелонефрита обеспечивается при $\bar{c}(t)=150\pm50$ мкг/мл (см. [95]). По этой функции $\bar{c}(t)=150$ и по экспериментально регистрируемой зависимости $K(t)$ с помощью алгоритма о.п.н. определялся,

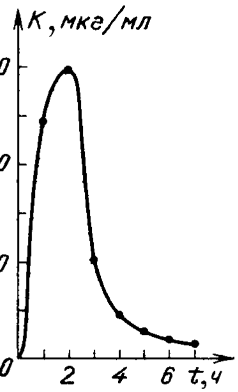

Рис. 27. Экспериментальная весовая функция

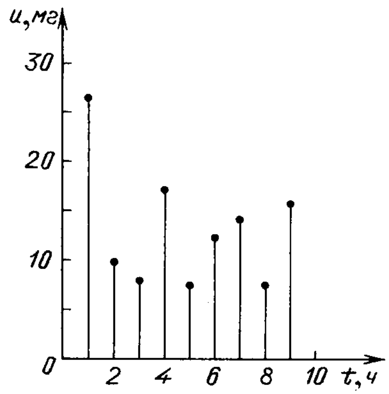

Рис. 28. Оптимизированный режим введения препарата

например, режим введения препарата, оптимальный в смысле минимума токсического воздействия на организм больного. Приведем конкретные результаты.

На рис. 27 показана экспериментальная зависимость $K(t)$ для одного из больных. На рис. 28 даны рекомендуемые дозы препарата, полученные в результате расчета на ЭВМ решения задачи (5), оптимального по функционалу Ω вида (7) (минимальное токсическое воздействие). Этот режим был реализован на практике. Наблюдаемая в результате этого экспериментальная зависимость $c(t)$, обусловленная такой дозировкой, показана на рис. 29. Видно, что почти все значения $c(t)$ входят в допустимую полосу разброса ± 50 мкг/мл около величины

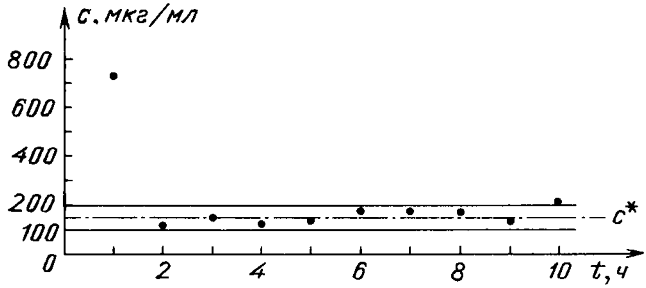

Рис. 29. Экспериментальная зависимость $c(t)$, полученная в результате применения оптимизированного режима лечения

$\bar{c}(t)$. Для сравнения на рис. 30 приведена экспериментальная зависимость $c(t)$, полученная у того же больного при реализации рекомендаций фирмы — изготовителя препарата. Видно, что концентрации $c(t)$ превосходят почти во всех точках оптимальную величину $\bar{c}(t)$ в 3—7 раз.

Применение оптимизации дозировки позволило снизить суточную дозу лекарства по сравнению с инструкцией фирмы-изготовителя

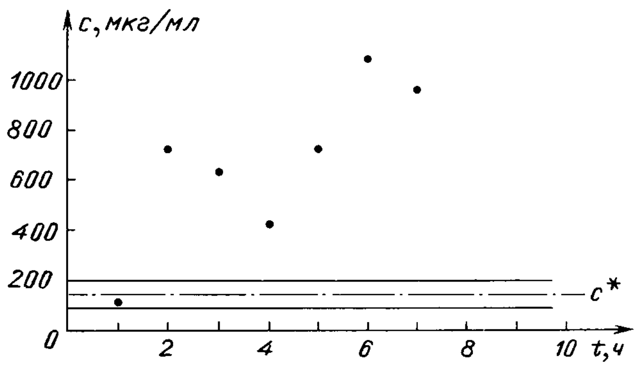

Рис. 30. Экспериментальная зависимость $c(t)$, полученная в результате применения к тому же больному режима лечения, рекомендованного фирмой — изготовителем препарата

в 3—5 раз и соответственно уменьшить степень токсического воздействия препарата на больного.

Проводились также расчеты по оптимизации времен введения фиксированных доз препарата.

§ 3. Оптимальное математическое проектирование высокотемпературных излучателей

В ряде технологических процессов, связанных с термической обработкой изделий, широко используются в качестве источников нагрева системы высокотемпературных излучателей, которые в дальнейшем будем кратко называть нагревателем. В связи с проектированием конкретных нагревателей возникает задача определения такого поля температур на его поверхности, которое обеспечило бы за счет излучения тепла требуемое поле температур, а также тепловых потоков на изделии. Ниже рассматривается некоторая математическая модель, которая позволяет в определенном приближении решить эту задачу.

Предположим, что заданы поверхности изделия F_1 и нагревателя F_2. Пусть для простоты поверхности F_1 гладкая, а F_2 плоская, причем они не пересекаются. На поверхностях F_1, F_2 считаем определенными соответственно температурные поля $T(M_1)$, $T(M_2)$ (M_i — точка на поверхности F_i ($i=1, 2$)). Поле $T(M_1)$ считается заданным, а $T(M_2)$ подлежит определению. Поверхности F_1, F_2 предполагаются диффузно и селективно излучающими и отражающими излучение [176].

При этом заданы зависимости излучательной способности $\varepsilon=\varepsilon_i(T, \lambda)$ на каждой поверхности F_i от температуры T и длины волны излучения λ. Тогда известна и отражательная способность $r_i(T, \lambda)=1-\varepsilon_i(T, \lambda)$, а в соответствие с законом Кирхгофа [176] и поглощательная способность $a_i(T, \lambda)=\varepsilon_i(T, \lambda)$. Излучение, падающее на систему F_1, F_2 извне, отсутствует.

Зададим функцию $q(M_1, \lambda)$ — поверхностную плотность результирующего излучения с поверхности F_1 (результирующий тепловой поток с F_1). Считая все упомянутые функции, характеризующие процесс обмена излучением между поверхностями F_1, F_2, непрерывными, можно написать интегральное уравнение, которое связывает заданные величины $T(M_1)$, $q(M_1, \lambda)$ с искомой функцией $T(M_2)$ (см. [174, 175]):

$$\iint\limits_{F_2} K(M_1, N_2)\varepsilon_\Sigma[T(N_2)][T(N_2)/T_0]^4\, dS_{N_2}+$$
$$+\iint\limits_{F_2}\iint\limits_{F_1} K(M_1, N_2)K(N_2, N_1)B[T(N_2), N_1]\, dS_{N_2}\, dS_{N_1}=u(M_1). \tag{1}$$

При этом предполагается выполненным естественное условие на искомую величину: $T(M_2)\geqslant T_{\min}>0$, где $T_{\min}$ — известная константа.

В уравнении (1) приняты следующие обозначения: $N_{1,2}$ — точки на $F_{1,2}$, dS_{N_i} — элемент площади поверхности F_i, $K(M_i, N_j)\,dS_{N_j}$ (i, $j=1, 2$) — так называемый элементарный угловой коэффициент площадки dS_{M_i} относительно площадки dS_{N_j} (см. [176]). Функции $K(M_i, N_j)$ определяются геометрией поверхностей $F_{1,2}$ и поэтому известны. Известна также функция $\varepsilon_\Sigma(T)$ — интегральная полусферическая излучательная способность поверхности F_2. Она определяется экспериментально и известна для многих материалов. Константа T_0 представляет собой известную среднюю температуру на поверхности изделия F_1. Оператор $B[T(N_2), N_1]$, действующий на функцию $T(N_2)$, определяется формулой

$$B[T(N_2), N_1]=\sigma_0^{-1} T_0^{-4} \int_0^\infty r_2 [T(N_2), \lambda] \{E_0 [T(N_1), \lambda]-$$

$$-r_1 [T(N_1), \lambda] \cdot a_1^{-1} [T(N_1), \lambda] \cdot q(N_1, \lambda)\} d\lambda,$$

в которой функция $E_0(T, \lambda)$ задается известной формулой Планка, а σ_0 — постоянная Стефана — Больцмана [176].

Функция $u(M_1)$ в правой части уравнения (1) может быть вычислена по заданным величинам K, $T(M_1)$, T_0, $q(M_1, \lambda)$, $r_1(T, \lambda)$, $a_1(T, \lambda)$:

$$u(M_1)=\left[\frac{T(M_1)}{T_0}\right]^4 - \iint_{F_1} K(M_1, N_1)\left[\frac{T(N_1)}{T_0}\right]^4 dS_{N_1} -$$

$$-\sigma_0^{-1} T_0^{-4} \int_0^\infty q(M_1, \lambda) a_1^{-1} [T(M_1), \lambda]\, d\lambda - \sigma_0^{-1} T_0^{-4} \times$$

$$\times \iint_{F_1} K(M_1, N_1) \left\{ \int_0^\infty r_1 [T(N_1), \lambda] \cdot a_1^{-1} [T(N_1), \lambda] q(N_1, \lambda) d\lambda \right\} dS_{N_1}.$$

Запишем нелинейное интегральное уравнение (1) для нахождения поля температур $T(N_2)$ на нагревателе в операторной форме:

$$A[M_1, T(N_2)]=u(M_1), \quad u(M_1)\in L_2(F_1), \quad T(N_2)\in D, \qquad (2)$$

где A — нелинейный интегральный оператор, действующий из множества решений $D=\{T(N_2)\in C(F_2): T(N_2)\geqslant T_{\min}\}$ в $L_2(F_1)$ и определяемый левой частью равенства (1). Операторное уравнение (2) относится к числу некорректно поставленных задач: оно может не иметь решений на множестве D в случае, если заданное поле $T(M_1)$ на изделии физически нереализуемо. Это уравнение может быть также неустойчивым по отношению к возмущениям оператора A, вызванным использованием экспериментальных величин ε_Σ, r_1, a_1, а также к неизбежным ошибкам, вызванным дискретизацией оператора из (2) при решении этого уравнения на ЭВМ. Поэтому естественно поставить задачу о поиске Ω-оптимальных квазирешений уравнения (2) на множестве D по схеме из § 2.1.

Зададим функционал

$$\Omega[T(N_2)]=\iint_{F_2} \varepsilon_\Sigma [T(N_2)]\left[\frac{T(N_2)}{T_0}\right]^4 dS_{N_2}. \qquad (3)$$

Этот функционал имеет физический смысл: он представляет собой значение полной собственной энергии, излучаемой с поверхности нагревателя [175]. Температурное поле $\bar{T}(N_2)$, являющееся Ω-оптимальным квазирешением уравнения (2) на множестве D для такого функционала Ω, не только наилучшим образом (в смысле минимума нормы в $L_2(F_1)$) удовлетворяет уравнению (1), (2), но и обеспечивает минимальные затраты энергии при работе такого нагревателя.

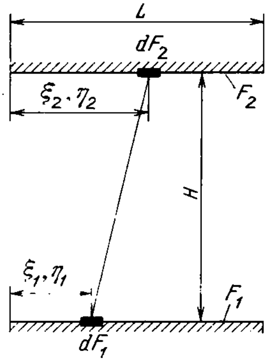

Рис. 31. Схема излучающей системы

Для нахождения Ω-оптимальных квазирешений уравнения (2) на множестве D можно использовать алгоритмы из гл. 2 и в том числе алгоритм обобщенного принципа невязки. Применялась программная реализация, описанная в § 1.14, 3.5, учитывающая требование $T(N_2) \geqslant T_{\min}$. Таким образом был решен ряд задач по проектированию оптимального распределения поля температур на нагревателе с целью создания заданных величин $T(M_1)$, $q(M_1, \lambda)$ на изделии. Приведем простейший пример.

Рассмотрим поверхности F_1, F_2, представляющие собой две параллельные бесконечные полосы ширины L, перпендикулярные плоскости чертежа (рис. 31) и расположенные на расстоянии H друг от друга. Материал полос — полированный вольфрам. Зависимость $\varepsilon(T, \lambda)$ для этого материала в диапазоне длин волн λ от 0,3 до 5 мкм хорошо известна. Функция $\varepsilon_\Sigma(T)$ для вольфрама также экспериментально найдена.

В данном примере рассматривалась задача о нахождении такого поля температур на нагревателе, которое обеспечило бы создание на изделии постоянной температуры $T(M_1) = T_1 = 1000$ К при условии, что $q(M, \lambda) = 0$. При этом предполагается выполненным ограничение на искомое решение $T(M_2) \geqslant T_{\min} = 300$ К. Уравнение (1) в рассматриваемом случае принимает вид

$$\int_0^1 K(\xi_1, \eta_2)\varepsilon_\Sigma[T(\eta_2)][T(\eta_2)/T_1]^4 d\eta_2 + \\ + \int_0^1 K(\xi_1, \eta_2)\varphi(\eta_2)B[T(\eta_2)]d\eta_2 = 1 \equiv u(\xi_1),$$

$$T(\eta_2)/T_1 \geqslant 0{,}3, \tag{4}$$

где

$$B[T(\eta_2)] = 1 - \frac{1}{\sigma_0 T_1^4}\int_0^\infty \varepsilon[T(\eta_2), \lambda]E_0(T_1, \lambda)d\lambda,$$

$$K(\xi, \eta) = \gamma^2[(\xi - \eta)^2 + \gamma^2]^{-3/2}/2, \quad \gamma \equiv H/L,$$

$$\varphi(\eta_2) = \int_0^1 K(\eta_2, \eta_1)d\eta_1.$$

Величины ξ_1, η_1, η_2 суть безразмерные (отнесенные к L) координаты на F_1, F_2. Функционал (3) использовался в виде

$$\Omega[T(\eta_2)] = \int_0^1 \varepsilon_\Sigma[T(\eta_2)] \left[\frac{T(\eta_2)}{T_1}\right]^4 d\eta_2.$$

На рис. 32 представлены результаты нахождения Ω-оптимального квазирешения уравнения (4) $\theta(\eta_2) \equiv \bar{T}(\eta_2)/T_1$ для различных значений γ относительного расстояния между поверхностями. Необходимо

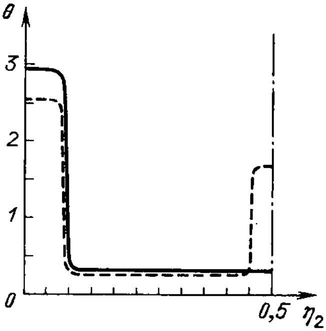

Рис. 32. Распределение относительной температуры $\theta(\eta_2)$ на поверхности нагревателя F_2 при различных γ: $\gamma=1$ —; $\gamma=0{,}5$ — — —

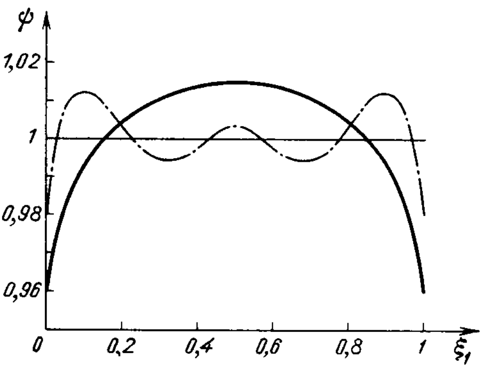

Рис. 33. Сравнение зависимостей $\psi(\xi_1)$, отвечающих найденным функциям $\theta(\eta_2)$ при различных γ: $\gamma=1$ —; $\gamma=0{,}5$ — · —; $u(\xi_1)\equiv 1$

отметить, что требование $T(M_1)=T_1$ в данной системе заведомо физически нереализуемо. Поэтому представляет интерес поведение невязки уравнения (4) на найденном Ω-оптимальном квазирешении $\bar{T}(\eta_2)$. На рис. 33 приведен график зависимости левой части уравнения (4) с $T(\eta_2)=\bar{T}(\eta_2)$ от ξ_1:

$$\psi(\xi_1) = \int_0^1 K(\xi_1, \eta_2)\varepsilon_\Sigma[\bar{T}(\eta_2)] \left[\frac{\bar{T}(\eta_2)}{T_1}\right]^4 d\eta_2 + \int_0^1 K(\xi_1, \eta_2)\varphi(\eta_2)B[\bar{T}(\eta_2)]\,d\eta_2$$

по сравнению с требуемой правой частью уравнения $u(\xi_1)=1$. Видно, что различие их составляет 2—4% в зависимости от рассматриваемого γ.

Приведенные расчеты, как и многие другие, показывают, что для обеспечения постоянства температуры $T(M_1)$ на изделии следует применять нагреватели с зонной структурой поля температур. При этом ширина зон, их положение на нагревателе и их температура определяются при расчетах по рассматриваемой методике.

§ 4. Восстановление локальных профилей линии поглощения элементов по наблюдаемым изменениям в спектре Ар-звезд

В соответствии с [66, 70, 72] рассмотрим следующую нелинейную обратную задачу астрофизики. Многие звезды, в том числе так называемые Ар-звезды, показывают периодические изменения профилей линий поглощения некоторых элементов, которые могут быть объяснены неравномерностью распределения химических элементов по поверхности вращающейся звезды, а следовательно, различием профилей линий поглощения, образующихся в различных участках звездной поверхности. Тот факт, что переменность профилей обусловливается именно неоднородностью химического состава, а не различием физических условий на поверхности звезды, подтверждается тем, что спектральный состав непрерывного излучения и яркость звезды, определяемые в основном физическими условиями в атмосфере звезды, не изменяются или изменяются весьма слабо. Кроме того, об этом говорят часто наблюдающаяся синхронность изменения профилей различных линий данного элемента, соответствующих различной степени его ионизации, и асинхронность изменений линий различных элементов, даже если они имеют близкие потенциалы возбуждения и ионизации.

Очень часто звезды, подобные описанным, обладают значительными магнитными полями, изменяющимися с теми же периодами, что и профили линий. Весьма вероятно, что неоднородность химического состава обусловлена наличием на звезде сильного магнитного поля. Для выяснения этого было бы особенно интересно локализовать аномалии химического состава с тем, чтобы сравнить их с распределением магнитного поля.

Предметом нашего рассмотрения является задача восстановления локальных профилей линий поглощения, т. е. определение по профилям линий поглощения, наблюдаемым в различные моменты времени, профилей линий, образующихся в фиксированной точке звездной поверхности. Дальнейшая обработка этих профилей классическими методами позволит получить число поглощающих атомов или ионов в каждой точке поверхности. Опишем подробнее постановку задачи [218]. Рассмотрим вопрос о профилях линий поглощения в спектре звезды, наблюдаемом на Земле. Для этого воспользуемся определением интенсивности излучения $I(M;\ \theta,\ \lambda)$ в точке M поверхности звезды на частоте $\nu=c/\lambda$ (c — скорость света) под углом θ к нормали через количество энергии dE, переносимой излучением через единичную площадку поверхности под углом θ к ее нормали в телесном угле $2\pi\sin\theta\, d\theta$ в отрезке длин волн $[\lambda,\ \lambda+d\lambda]$ интервале времени $[t,\ t+dt]$ посредством равенства

$$dE=2\pi I(M;\ \theta,\ \lambda)\sin\theta\, d\theta\cos\theta\, d\lambda\, dt. \tag{1}$$

Пользуясь таким образом определенной величиной интенсивности излучения, мы получаем, что энергия, приходящая на Землю от всей видимой поверхности звезды, пропорциональна величине

$$F_\lambda^+=\iint I(M;\ \theta,\ \lambda)\cos\theta\, dM, \tag{2}$$

где dM — элемент сферической поверхности, θ — угол между нормалью к поверхности звезды в точке M и лучом зрения; интегрирование производится по всей видимой поверхности звезды. Если в качестве системы координат на сфере взять сферическую систему координат (θ, φ) с осью, направленной по лучу зрения, то $dM = \sin\theta\, d\theta\, d\varphi$ и видимая поверхность определяется условием $\cos\theta > 0$ (рис. 34).

Для достаточно быстро вращающихся звезд, какими являются Ар-звезды, следует учитывать, что доплеровское смещение излучаемого

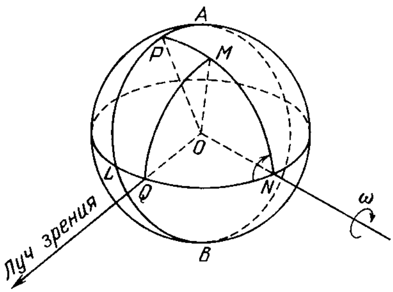

Рис. 34. Система координат: $ABPA$ — экватор, $LQNM = l - \omega t$, $LPOM = \Phi$ широта, $LNQM = \varphi$, $LQOM = \theta$

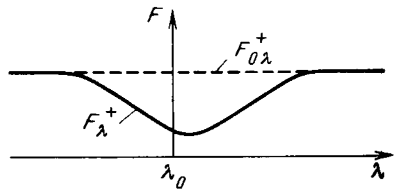

Рис. 35. График функции $F(\lambda)$

света в различных точках поверхности звезды будет различным и может достигать значений, сравнимых с шириной линии. Тем самым под интегралом в выражении, определяющем F^+_λ, в функции $I(M; \theta, \lambda)$ аргумент λ должен быть заменен на $\lambda + \Delta\lambda_D(M)$, где $\Delta\lambda_D(M)$ описывает величину доплеровского смещения излучения, образовавшегося в точке M поверхности вращающейся звезды.

Будем различать интенсивность излучения в непрерывном спектре, т. е. вдали от центра линии (или интенсивность излучения при отсутствии поглощающих атомов) $I_0(M; \theta, \lambda)$, и интенсивность излучения в линии, которая естественно зависит от содержания поглощающих атомов в атмосфере.

Спектр, который мы регистрируем на Земле, условно изображен на рис. 35. Здесь $F^+_{0\lambda}$ — поток, обусловленный непрерывным спектром, а F^+_λ — поток, обусловленный наличием поглощения в линии. Поскольку поглощение в земной атмосфере одинаково для спектра линии и непрерывного излучения, то для того, чтобы освободиться от влияния этого поглощения, обычно рассматривается величина

$$R(\lambda) = 1 - \frac{F^+_\lambda}{F^+_{0\lambda}} = 1 - \frac{\iint\limits_{\cos\theta>0} I(M; \theta, \lambda + \Delta\lambda_D(M)) \cos\theta\, dM}{\iint\limits_{\cos\theta>0} I_0(M; \theta, \lambda + \Delta\lambda_D(M)) \cos\theta\, dM}, \tag{3}$$

которая уже не зависит от состояния земной атмосферы и называется *профилем линии*.

Функции $I_0(M; \theta, \lambda)$, $I(M; \theta, \lambda)$ являются решениями уравнений переноса излучения (см., например, [213]), в которые входят коэффициенты поглощения и переизлучения в линии, а также коэффициенты поглощения и излучения в непрерывном спектре как функции глубины в атмосфере. Коэффициенты в непрерывном спектре определяются в основном зависимостью концентрации водорода и его ионов от глубины, или, как это обычно называют, *моделью атмосферы*.

Поскольку ширины рассматриваемых линий обычно малы, то в пределах линии интенсивность излучения в непрерывном спектре практически не изменяется, и мы будем считать ее не зависящей от λ. Кроме того, предположим, что интенсивность в непрерывном спектре не зависит от точки M на поверхности звезды, так как физические условия на звезде приблизительно однородны. В соответствии со сделанными замечаниями $I_0(M; \theta, \lambda) = I_0(\theta)$. Зависимость интенсивности излучения в непрерывном спектре от угла (так называемое *потемнение к краю*) будем рассматривать в виде $I_0(\theta) = I_0 \cdot u_1(\theta)$, где $u_1(\theta)$ — линейная функция $\cos\theta$ (см., например, [5]). Такой закон потемнения к краю, полученный из теоретических соображений, хорошо согласуется с наблюдаемыми изменениями интенсивности излучения от центра к краю солнечного диска, а также данными, полученными при интерпретации кривых блеска двойных затменных систем. Зависимость коэффициента u_1 в выражении для функции $u_1(\theta) = 1 - u_1 + u_1 \cos\theta$ от температуры хорошо изучена, а температура звезды может быть определена по ее спектральному классу.

Как показано в [218], функция

$$r(M, \lambda) = 1 - \frac{I(M; 0, \lambda)}{I_0(M; 0)} \tag{4}$$

во всех случаях хорошо представляется формулой Миннаэрта

$$r(M, \lambda) = \left\{ \frac{1}{x(M, \lambda)} + \frac{1}{R_c(M)} \right\}^{-1}, \tag{5}$$

где $R_c(M)$ — центральная глубина линии, $x(M, \lambda) = \tau(M) k(\lambda)$; $k(\lambda)$ — так называемый *фойгтовский профиль* (см., например, [213]) — известная функция от λ. Коэффициент $\tau(M)$, зависящий от точки на поверхности звезды, определяется относительным содержанием поглощающего элемента в атмосфере.

Зависимость интенсивности излучения в линии $I(M; \theta, \lambda)$ от угла θ может быть очень сложной. Линии поглощения могут к краю звездного диска и усиливаться, и даже превращаться в эмиссионные. Возможен случай, когда центральные части линии (ядро линии) ослабляются к краю, а крылья усиливаются. Однако для многих линий эти изменения от центра к краю незначительны, и мы будем в соответствии с работой [218] предполагать, что эту зависимость можно записать в виде

$$1 - \frac{I(M; \theta, \lambda)}{I_0(M; \theta)} = r(M, \lambda) u_2(\theta), \tag{6}$$

где функция $u_2(\theta)$ линейна относительно $\cos\theta$: $u_2(\theta)=1-u_2+u_2\cos\theta$. Параметр u_2 этой функции можно найти, например, путем численного интегрирования уравнений переноса для модели атмосферы, близкой к атмосфере рассматриваемой звезды. Для слабо меняющихся от центра к краю профилей коэффициент принимается равным нулю.

В условиях однородности физических условий параметры τ и R_c формулы Миннаэрта (5) не являются независимыми. Действительно, величина R_c определяет центральную глубину линии, а τ — ее ширину. Таким образом, при малых R_c и больших τ формула (5) описывает мелкую и широкую линию и, наоборот, при больших R_c и малых τ — глубокую и узкую линию. При фиксированной температуре и других физических параметрах атмосферы такие профили не могут существовать одновременно. С увеличением концентрации поглощающего элемента линия становится глубже, пока в центре линии не наступит насыщение, после чего линия только расширяется. Центральная глубина при насыщении является при этом функцией поверхностной температуры.

Как показано в [169] при изучении полученных при численном решении уравнений переноса профилей линий, параметры R_c, τ являются зависимыми, и зависимость центральной глубины от величины τ можно аппроксимировать формулой

$$R_c=h(\tau)=c_1\left[1-\exp\{-c_2\tau\}\right]. \tag{7}$$

Параметры c_1, c_2 должны быть подобраны для каждой звезды и линии на основании исследования модели атмосферы, наиболее пригодной для этой звезды.

В окончательном виде наблюдаемые на Земле профили линии поглощения в спектре интегрального излучения всего диска звезды определяются на поверхности звезды функцией $\tau(M)$ по формуле

$$R(\lambda)=\frac{1}{N}\iint\limits_{\cos\theta>0} f\left[\tau(M),\quad \lambda+\Delta\lambda_D(M)\right]u_1(\theta)u_2(\theta)\cos\theta\,dM, \tag{8}$$

где

$$N=2\pi\int\limits_0^{\pi/2}\cos\theta u_1(\theta)\sin\theta\,d\theta, \tag{9}$$

θ — угол между нормалью в точке M и лучом зрения,

$$f(\tau,\ \lambda)=\left\{\frac{1}{h(\tau)}+\frac{1}{\tau k(\lambda)}\right\}^{-1}, \tag{10}$$

$\Delta\lambda_D(M)$ — доплеровское смещение излучения за счет вращения звезды.

Заметим теперь, что при вращении звезды профили $R(\lambda)$ будут изменяться. В соответствии с этим будем писать $R(\lambda,\ \omega t)$, где t — время, ω — угловая скорость вращения звезды, и уравнение (8) запишем в виде

$$R(\lambda,\ \omega t)=\iint f\left[\tau(M),\ \lambda+\Delta\lambda_D(M,\quad \omega t)\right]H(M,\ \omega t)\,dM. \tag{11}$$

Если в качестве координат на поверхности звезды выбрать широту Φ и долготу l (см. рис. 34), то элемент площади поверхности звезды есть $dM=\cos\Phi\, d\Phi\, dl$, функция $f(\tau,\ \lambda)$ определена выше,

$$\Delta\lambda_D(\Phi,\ l,\ \omega t)=v_{\text{экв}}\sin i\cos\Phi\cos(l-\omega t)\lambda_0/c, \tag{12}$$

$$H(\Phi,\ l,\ \omega t)=\left\{\frac{1}{N}u_1(\theta)u_2(\theta)\cos\theta\cos\Phi:\ \cos\theta>0;\ 0:\ \cos\theta\leqslant 0\right\}, \tag{13}$$

$$\cos\theta=\cos i\sin\Phi+\cos\Phi\sin i\cos(l-\omega t) \tag{14}$$

(здесь $v_{\text{экв}}$ — экваториальная скорость вращения звезды, а i — угол наклона оси вращения звезды к лучу зрения).

Фойгтовский профиль $k(\lambda)$, определяющий зависимость профиля от длины волны, является сверткой двух функций, описывающих профиль линии, уширенной за счет давления (столкновений) и доплеровского смещения при микротурбулентных движениях в атмосфере звезды:

$$k(\lambda)=\frac{\varkappa_l}{\varkappa_d}\frac{1}{\pi}\int\limits_{-\infty}^{+\infty}\frac{e^{-t^2}dt}{\varkappa_l^2/\varkappa_d^2+(t-\lambda/\varkappa_d)^2}. \tag{15}$$

Здесь $\varkappa_l$ — величина лоренцевского уширения давлением. Параметр $\varkappa_d$ носит во многом формальный характер и характеризует величину уширения линии доплеровским смещением. Этот параметр должен подбираться на основе анализа подходящей модели атмосферы таким образом, чтобы формула Миннаэрта хорошо описывала профили линий поглощения при различных концентрациях поглощающих атомов.

Мы будем ставить своей задачей нахождение функции $\tau(M)$ по измеренным в различные фазы профилям линии — функции $R(\lambda,\ \omega t)$. При этом соотношение (11) мы будем рассматривать как нелинейное интегральное уравнение для определения функции $\tau(M)$. На функцию $\tau(M)$ при этом должно быть наложено естественное ограничение $\tau(M)\geqslant 0$.

Знание функции $\tau(M)$ означает фактически знание локальных профилей линий поглощения $r(M,\ \lambda)$, а следовательно, и содержания в точке M поверхности звезды элемента, ответственного за образование рассматриваемой линии поглощения.

Остается сделать замечание относительно отыскания некоторых параметров, входящих в уравнение (11), которые должны быть предварительно определены из независимых наблюдений. Это касается в первую очередь параметров $v_{\text{экв}}$, $\sin i$. Величину $v_{\text{экв}}\sin i$ можно оценить, измерив полуширину $\Delta\lambda_{0,5}$ некоторых неизменяющихся линий, уширение которых связано в основном с вращением звезды. Тогда $v_{\text{экв}}\sin i\approx 1{,}25c\Delta\lambda_{0,5}/(2\lambda)$, где c — скорость света. Кроме того, обычно можно считать, что радиус R для звезд этого класса равен 2—4 $R_\odot$ ($R_\odot$ — радиус Солнца), и оценить его, исходя из спектрального класса звезды. Тогда $v_{\text{экв}}=2\pi R/P$, где P — период вращения звезды. Постановка задачи в такой форме впервые дана в [69].

Введем некоторые полезные для дальнейшего определения. Будем называть *эквивалентной шириной линии* величину, быть может зависящую от времени,

$$W(\omega t)=\int\limits_0^{\infty}R(\lambda,\ \omega t)\,d\lambda=\iint\xi(M)H(M,\ \omega t)\,dM, \tag{16}$$

где

$$\xi(M)=\int_0^\infty r(M,\ \lambda)\,d\lambda \tag{17}$$

есть локальная эквивалентная ширина линии. Определим также *эффективное доплеровское смещение* как

$$Z(\omega t)=\int_0^\infty R(\lambda,\ \omega t)\,(\lambda-\lambda_0)\,d\lambda=\iint\xi(M)\Delta\lambda_D(M,\ \omega t)H(M,\ \omega t)\,dM, \tag{18}$$

где λ_0 — центр рассматриваемой линии.

Рассмотрим сначала методы исследования химического состава нормальных звезд с однородным распределением поглощающего вещества на поверхности, а следовательно, с $r(M,\ \lambda)$, $\xi(M)$, не зависящими от M. В этом случае обычно используются только наблюдаемые эквивалентные ширины W. При этом на основании исследования распределения энергии в непрерывном спектре излучения звезды и его водородных и гелиевых линий поглощения находятся параметры атмосферы: температура $T_{\text{эфф}}$, ускорение силы тяжести g, отношение содержания гелия и водорода. На основании значений этих параметров может быть рассчитана модель атмосферы, и на основании этой модели, решая уравнение переноса, можно рассчитать зависимость эквивалентной ширины линии поглощения от относительного содержания поглощающего элемента в атмосфере (так называемую кривую роста). По полученной кривой роста уже можно определить обилие поглощающего вещества на основе знания эквивалентных ширин его линий поглощения. Метод определения химического состава, основанный на расчетах моделей атмосфер, принято называть методом моделей атмосфер. Эти методы успешно применялись для решения одной из важнейших проблем астрофизики — проблемы распространенности химических элементов, подробно рассмотренной в [5].

Для звезд с неоднородным распределением химического состава по поверхности задача отыскания распределения химического вещества, ответственного за образование данной линии, на поверхности звезды распадается на следующие задачи: а) задачу определения локальных эквивалентных ширин или профилей линий, б) задачу определения по этим эквивалентным ширинам или профилям линий локального химического состава.

Задача б) решается описанным выше методом моделей атмосфер. Перейдем теперь к вопросам, связанным с применением метода регуляризации некорректно поставленных задач для решения обратной задачи восстановления локальных профилей линий поглощения, т. е. для решения задачи а).

Для уравнения (11) запишем сглаживающий функционал

$$M^\alpha[\tau]=\alpha\iint\left\{\left(\frac{\partial\tau}{\partial\Phi}\right)^2+\left(\frac{\partial\tau}{\partial l}\right)^2\right\}d\Phi\,dl+\iint\{R(\lambda,\ \omega t)-$$
$$-\iint f[\tau(M),\ \lambda+\Delta\lambda_D(M,\ \omega t)]\cdot H(M,\ \omega t)\,dM\}^2\,d\lambda\,d\omega t. \tag{21}$$

Будем предполагать, что функция $\tau(M)$ принадлежит пространству функции W_2^1 на поверхности звезды, причем минимизацию функционала (21) необходимо производить на множестве функций $\tau(M)\geqslant 0$. Выбор параметра регуляризации α будем производить по альтернативному принципу (см. § 2.5).

Функцию $\tau(M)$ будем искать на множестве $\Phi\in[-i, \pi/2]$, $l\in\in[-\pi, \pi]$, поскольку участки звезды с $\Phi\in[-\pi/2, -i]$ не видны с Земли. Функцию $R(\lambda, \omega t)$, описывающую экспериментальные данные, будем считать заданной при $\lambda\in(\lambda_0-\Delta\lambda/2, \lambda_0+\Delta\lambda/2)$, $l\in[-\pi, \pi]$. Таким образом, в выражении (21) для $M^\alpha[\tau]$ интегрирование необходимо производить по множеству $\Phi\in[-i, \pi/2]$, $l\in[-\pi, \pi]$.

Перейдем теперь к конечноразностной аппроксимации задачи. На поверхности звезды в координатах $\Phi\in[-i, \pi/2]$, $l\in[-\pi, \pi]$ зададим прямоугольную сетку $\{\varphi_p, l_q\}$ из $N\times M$ узлов. Будем считать, что каждый узел сетки является центром ячейки на поверхности звезды, причем в каждой ячейке значение функции $\tau(\Phi, l)$ постоянно: $\tau(\Phi, l)=\tau_{pq}$ $(p=1, ..., N,\ q=1, ..., M)$. По λ выберем равномерную сетку из N_λ точек, включающую в себя центр линии; обозначим узлы этой сетки λ_k $(k=1, ..., N_\lambda)$. Будем также считать, что профили наблюдаются в $N_{\omega t}$ фазах $(\omega t)_l$ $(l=1, ..., N_{\omega t})$. Интегралы по λ, ωt в выражении (21) будем аппроксимировать по формуле прямоугольников на описанных сетках, при этом будем обозначать $R_{kl}=R[\lambda_k,(\omega t)_l]$.

Вместо переменной l введем переменную $\psi=l-\omega t$. Интегрирование по ψ в (21) так же, как и интегрирование по l будет производиться от $-\pi$ до π. Но, согласно (13), (14), имеем $H(\Phi, \psi)=0$, если $\cos\theta=\sin\Phi\times\cos i+\cos\Phi\sin i\cos\psi\leqslant 0$, т. е. $H(\Phi, \psi)\neq 0$ только, если $-\Gamma(\Phi)\leqslant\psi\leqslant\Gamma(\Phi)$, где $\Gamma(\Phi)=\arccos(\operatorname{tg}\Phi\cdot\operatorname{ctg} i)$. Выберем по Φ, ψ равномерные сетки $\{\Phi_i\}$, $\{\psi_j\}$ с шагами h_1 h_2 соответственно. Обозначим через $L(\Phi)$ целую часть $[\Gamma(\Phi)/h_2]$. Пусть $p(i, j, l)$, $q(i, j, l)$ — координаты центра ячейки на поверхности звезды (из прямоугольника $\Phi\in[-i, \pi/2]$, $l\in[-\pi, \pi]$), которой принадлежит точка $(\Phi_i,\ \psi_j+(\omega t)_l)$.

Согласно формуле (15), фойгтовский профиль $k(\lambda)$ является сверткой двух функций. Вычисление этой свертки является трудоемкой операцией, проводить которую каждый раз нецелесообразно. В связи с этим для вычисления $k(\lambda)$ будем использовать следующий алгоритм. Заметим, что в интересующих нас диапазонах изменения параметров $\varkappa_l$, $\varkappa_d$ функция $k(\lambda)$ хорошо аппроксимируется по формуле из [146]:

$$k(\lambda)=\frac{\sqrt{\ln 2}}{\sqrt{\pi}\varkappa_v}(1-\xi)\exp\{-\eta^2\ln 2\}+\frac{1}{\pi\varkappa_v}\,\frac{1}{1+\eta^2}-$$

$$-\frac{\xi(1-\xi)}{\pi\varkappa_v}\left(\frac{1{,}5}{\ln 2}+1+\xi\right)\cdot\left[0{,}066\exp\{-0{,}4\eta^2\}-\frac{1}{40-5{,}5\eta^2+\eta^4}\right]; \tag{22}$$

$$\varkappa_v=0{,}5(\varkappa_l+\sqrt{\varkappa_l^2+4\varkappa_d^2})+0{,}05\varkappa_l\left(1-\frac{2\varkappa_l}{\varkappa_l+\sqrt{\varkappa_l^2+\varkappa_d^2}}\right);$$

$$\eta\equiv(\lambda-\lambda_0)/\varkappa_v,\quad \xi\equiv\varkappa_l/\varkappa_v.$$

Пользуясь формулой (22) составим таблицу значений $k(\lambda)$ на равномерной сетке по λ примерно из 300—600 точек так, чтобы она охватывала все значения λ, которые могут встретиться при вычислении (21). Значение $k(\lambda)$ в произвольной точке будет находиться линейной интерполяцией.

Окончательно конечноразностная аппроксимация функционала (21) имеет вид

$$\hat{M}^\alpha[\tau]=\alpha\hat{\Omega}[\tau]+\hat{\Phi}[\tau], \tag{23}$$

где

$$\hat{\Phi}[\tau]=\sum_{k=1}^{N_\lambda}\sum_{l=1}^{N_{\omega t}} w_{kl}\Bigg\{R_{kl}-$$

$$-\sum_{i=1}^{N}\sum_{j=-L(\Phi_i)}^{L(\Phi_i)} h_1 h_2 H(\Phi_i,\psi_j)\cdot f\left[\tau_{p(i,j,l)q(i,j,l)},\ \lambda_k+\Delta\lambda_D(\Phi_i,\psi_j)\right]\Bigg\}^2, \tag{24}$$

$$\hat{\Omega}(\tau)=\sum_{j=1}^{M}\sum_{i=2}^{N}(\tau_{ij}-\tau_{i-1j})^2+\sum_{i=1}^{N}\left\{\sum_{j=2}^{M}(\tau_{ij}-\tau_{ij-1})^2+(\tau_{i1}-\tau_{iM})^2\right\}. \tag{25}$$

Здесь w_{kl} — веса, введение которых мотивируется тем, что: точность измерений $R(\lambda,\omega t)$ и число измерений $R(\lambda,\omega t)$ в разных фазах различно; профили, соответствующие каждой фазе обычно известны с большей точностью в центре линии и с меньшей на краях линии (особенно из-за эффекта блендирования).

Таким образом, задача восстановления локальных профилей линий поглощения сводится к минимизации функционала $\hat{M}^\alpha[\tau]$ в конечномерном пространстве размерности $N\times M$ при ограничениях $\tau_{pq}\geqslant 0$ с последующим выбором параметра регуляризации по альтернативному принципу.

Для минимизации $\hat{M}^\alpha[\tau]$ предлагается использовать метод проекции сопряженных градиентов (см., например, [31, 91, 173]). Итерационный процесс строится следующим образом. Задается начальное приближение — вектор $\tau^{(0)}$ с неотрицательными компонентами. По заданной точке $\tau^{(i)}$ минимизирующей последовательности вычисляется градиент $g^{(i)}=\operatorname{grad}\hat{M}^\alpha[\tau^{(i)}]$ (для вычисления градиента $\hat{M}^\alpha[\tau]$ легко могут быть получены аналитические выражения). По направлению градиента на предыдущем шаге $g^{(i-1)}$ и направлению спуска на предыдущем шаге $h^{(i-1)}$ вычисляется новое направление спуска по формуле

$$h^{(i)}=g^{(i)}+\gamma_i h^{(i-1)}, \tag{26}$$

где

$$\gamma_i=\frac{(g^{(i)}-g^{(i-1)},\ g^{(i)})}{(g^{(i-1)},\ g^{(i-1)})} \tag{27}$$

(вариант Полака — Рибьера метода сопряженных градиентов [170]). После этого строится однопараметрическое множество, состоящее из элементов

$$\tau_\lambda=P(\tau^{(i)}-\lambda h^{(i)}),\quad \lambda\geqslant 0, \tag{28}$$

где P — оператор проектирования $N\times M$-мерного вектора на первый октант, и решается задача одномерной минимизации $\hat{M}^\alpha[\tau]$ на этом множестве. Точка минимума функции $\hat{M}^\alpha[\tau]$ на этом множестве принимается за следующий элемент $\tau^{(i+1)}$ минимизирующей последовательности. При решении одномерной задачи минимизации использовалась квадратичная аппроксимация функции

$$\theta(\lambda)\equiv\hat{M}^\alpha\left[P\left(\tau^{(i)}-\lambda h^{(i)}\right)\right]-\hat{M}^\alpha\left[\tau^{(i)}\right] \tag{29}$$

по трем точкам 0, $\lambda_\text{н}$, $2\lambda_\text{н}$. После этого приближенное решение одномерной задачи минимизации выбиралось в виде

$$\lambda_{\min}=\lambda_\text{н}\cdot\frac{1}{2}\,\frac{\theta_2-4\theta_1}{\theta_2-2\theta_1};\quad \theta_1\equiv\theta(\lambda_\text{н}),\quad \theta_2\equiv\theta(2\lambda_\text{н}). \tag{30}$$

При необходимости производится дополнительное уточнение этой точки. Через каждые $N\times M$ шагов минимизации производится «обновление» метода, т. е. γ полагается равным нулю в формуле (26). Кроме того, алгоритм обновляется, если вследствие неточности решения одномерной задачи минимизации на очередном шаге получится $(h^{(i)}, g^{(i)})\leqslant 0$.

Данный алгоритм был реализован на ЭВМ БЭСМ-6 в виде комплекса программ на языке Фортран и автокоде В. В. Степановым. Им же были проведены расчеты на ЭВМ приводимых ниже модельных задач и была обработана экспериментальная информация, любезно предоставленная В. Л. Хохловой (Астросовет АН СССР). В результате проведения модельных расчетов было исследовано влияние на профили наблюдаемых спектральных линий различных параметров, указанных в постановке задачи, а именно, коэффициентов потемнения к краю в непрерывном спектре и линии u_1, u_2, угла наклона оси вращения звезды к лучу зрения i, экваториальной скорости вращения $v_\text{экв}$, параметров фойгтовского профиля $\varkappa_l$, $\varkappa_d$, параметров зависимости $h(\tau)-c_1$, c_2. Было показано, что влияние параметров u_1, u_2, $\varkappa_l$, $\varkappa_d$, c_1, c_2 на интегральные профили линий невелико, и, следовательно, для решения обратной задачи достаточно знать хотя бы грубые оценки этих параметров, что в большинстве случаев удается получить из анализа экспериментальных данных, полученных от звезды.

Наиболее значительное влияние на наблюдаемые интегральные профили оказывают изменения параметров i, $v_\text{экв}$. Как показали расчеты, изменение параметра $v_\text{экв}$ изменяет профили линий $R(\lambda, \omega t)$ не только количественно, но и качественно. Это указывает на необходимость как можно более точного оценивания параметров i, $v_\text{экв}$ из независимых наблюдений.

Как уже отмечалось выше, область определения $\tau(M)$ различна при различных i. При этом область, видимая с Земли, тем больше, чем больше угол i. При $i=\pi/2$ видна вся поверхность звезды, однако при этом нарушается однозначность определения $\tau(M)$. При i, близких к нулю, доплеровское смещение излучения точек звездной поверхности будет мало и при этом восстановление распределения $\tau(M)$ окажется не возможным. Отсюда же ясно, что нельзя добиться хорошего

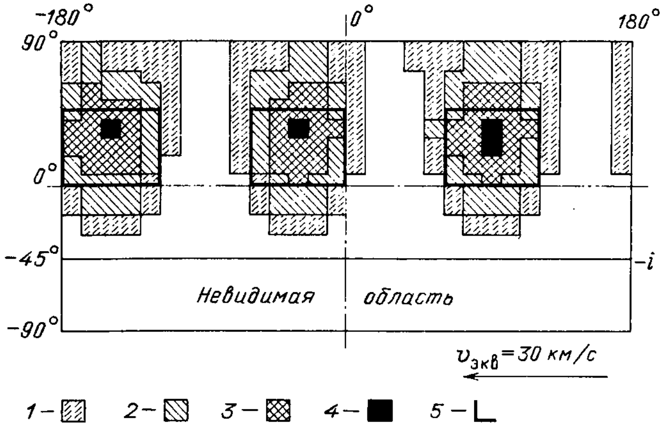

Рис. 36. Карта распределения локальных эквивалентных ширин (модель 1): *1* – 0,03 – 0,08 Å; *2* — 0,08 – 0,13 Å; *3* – 0,13 – 0,18 Å; *4* 0,18 – 0,23 Å; *5* контуры исходного распределения

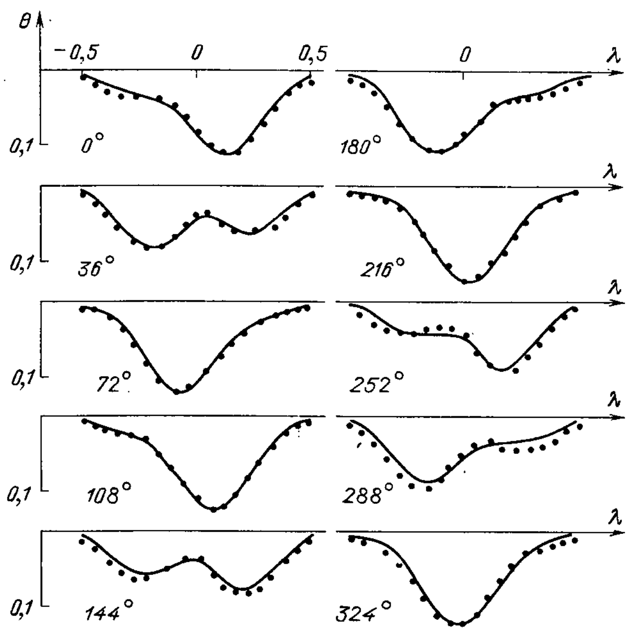

Рис. 37. Профили линий: модельные · · · ; вычисленные —. Среднеквадратичная погрешность – 0,005

восстановления $\tau(M)$ на границе области определения и в приполярных областях.

Приведем результаты расчетов некоторых модельных задач. Во всех случаях в качестве начального приближения при построении последовательности, минимизирующей сглаживающий функционал, использовалось однородное распределение $\tau(M)=0{,}7$. Этому значению параметра соответствует очень малое значение концентрации поглощающих элементов, при которой линия практически не образуется.

Поскольку при дальнейшей обработке локальных профилей в первую очередь представляет интерес распределение локальных эквивалентных ширин $\xi(M)$, то результаты расчетов мы будем представлять в терминах локальной эквивалентной ширины.

На рис. 36 изображена карта распределения локальных эквивалентных ширин, полученная в результате применения описанного метода к профилям линий, соответствующим наличию на поверхности звезды трех областей прямоугольной формы с повышенным содержанием поглощающего элемента (контуры этих областей — точное решение — также изображены на рис. 36). Как видно из рисунка положение пятен по долготе восстановилось достаточно хорошо. Размеры пятен по широте получились увеличенными, при этом интенсивность линии в пятнах несколько занижена по основанию с исходной. Следует отметить, что полученное распределение позволяет описывать «наблюдаемые профили» линий с хорошей точностью до 0,01, что соответствует значениям ошибок задания экспериментальной информации. Профили изображены на рис. 37.

Отметим, что изменения величины локальной эквивалентной ширины на поверхности соответствуют изменению относительного содержания поглощающего элемента на несколько порядков. Например, для линии λ 3856Å Si II изменения ξ от 0,1 Å до 0,2 Å соответствует изменению обилия Si примерно в 10 раз.

Результаты решения описанной модельной задачи демонстрируют возможности разделения областей с повышенным значением локальной эквивалентной ширины. Пятна, разделенные промежутком в 50° и более, достаточно надежно разрешаются.

Перечислим значения использованных параметров: $v_{\text{экв}}=35$ км/с, $i=50°$, $\varkappa_l=0{,}1989\cdot10^{-3}$, $\varkappa_d=0{,}0468$, $c_1=0{,}7$, $c_2=2{,}75\cdot10^{-3}$, $\lambda_0=3856$ Å, ширина линии $\Delta\lambda=1{,}5$ Å. Использовалась сетка по длине волны из 19 точек. Исследовались профили, соответствующие 10 фазам, равномерно распределенным по периоду. Число областей постоянства функции при решении обратной задачи $11\times30=330$. В модельном распределении значение локальной эквивалентной ширины в пятнах 0,2 Å. Вне пятен линия не образуется. Параметр регуляризации, найденный по альтернативному принципу, оказался $\alpha\sim10^{-7}$.

Рассмотрим теперь модельную задачу восстановления распределения локальных эквивалентных ширин $\xi(M)$, отличных от нуля в области, границы которой изображены на рис. 38. Во всей этой области значение τ постоянно и равно $1\cdot10^3$. На том же рисунке изображена карта распределения $\xi(M)$ по поверхности, полученная описываемым методом. Заметим, что в исходном распределении

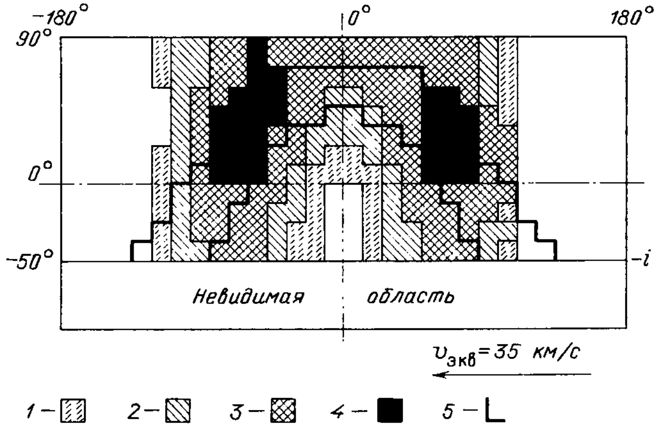

Рис. 38. Карта распределения локальных эквивалентных ширин (модель 2): *1* 0,02–0,05 Å; *2* – 0,05–0,08 Å; *3* 0,08–0,011 Å; *4* – 0,11–0,14 Å; *5* контуры исходного распределения

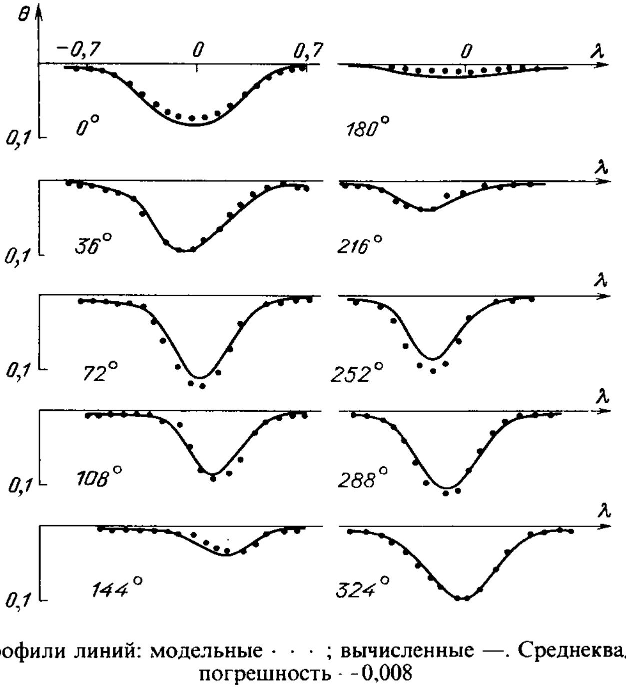

Рис. 39. Профили линий: модельные · · · ; вычисленные —. Среднеквадратичная погрешность – 0,008

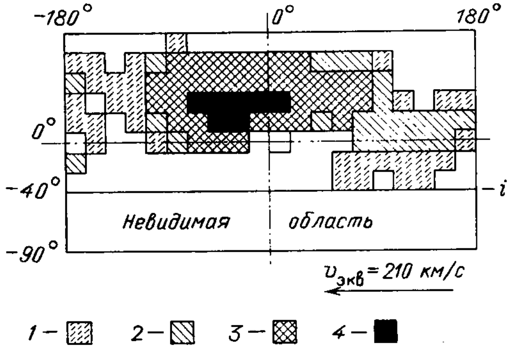

Рис. 40. HD 124224, λ 3862 Å Si II. Карта распределения локальных эквивалентных ширин: *1* — 0,1 – 0,2 Å; *2* — 0,2 – 0,3 Å; *3* — 0,3 – 0,4 Å; *4* — 0,4 – 0,5 Å

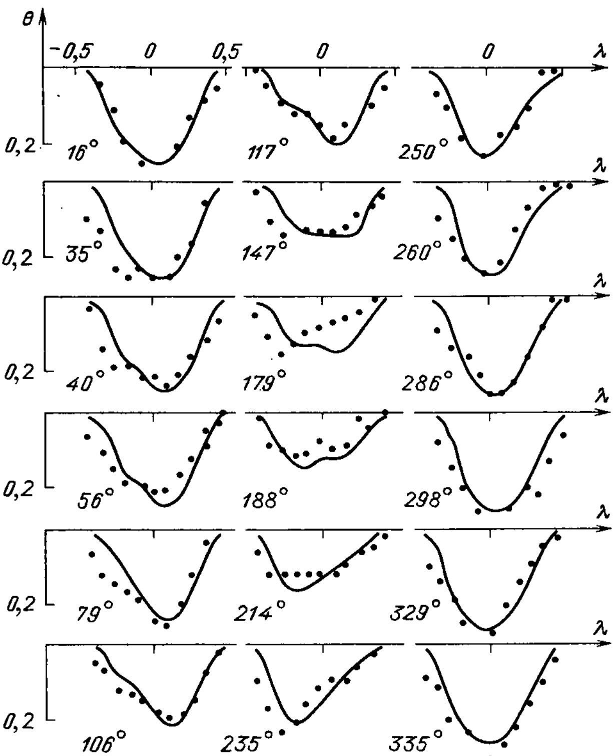

Рис. 41. HD 124224, λ 3862 Å Si II. Профили линий: наблюдаемые · · · ; вычисленные

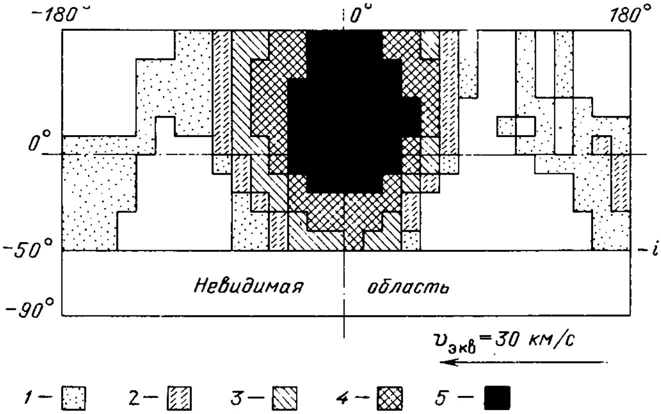

Рис. 42. HD 112412, λ 4129 Å Eu II. Карта распределения локальных эквивалентных ширин: *1* — 0,35 – 0,70 Å; *2* — 0,70 – 0,105 Å; *3* — 0,105 – 0,140 Å; *4* — 0,140 – 0,175 Å; *5* — 0,175 – 0,210 Å

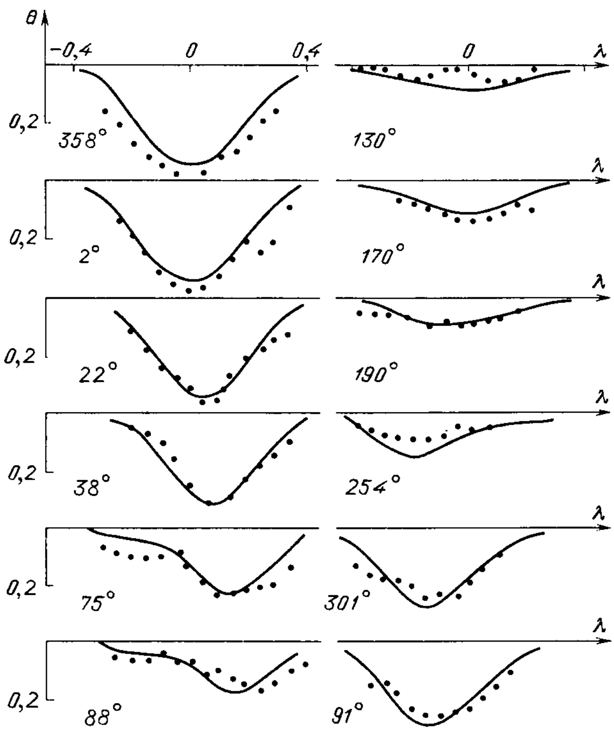

Рис. 43. HD 112412, λ 4129 Å Eu II. Профили линий: наблюдаемые · · · ; вычисленные —

область образования линии представляет собой кольцо большого радиуса. Эта модель демонстрирует возможности восстановления распределения $\tau(M)$ в различных областях звездной поверхности. В этом случае, так же, как и в предыдущем, профили линий, изображенные на рис. 39, приближаются очень хорошо. Значение параметра регуляризации $\alpha \sim 10^{-9}$. Остальные параметры те же, что и в предыдущей модельной задаче.

С помощью алгоритма, разработанного в данном параграфе, были обработаны профили спектральной линии λ 3862 Å Si II в спектре Ар-звезды CU Vir (HD 124224), полученные Т. А. Рябчиковой (Астросовет АН СССР) в Крымской астрофизической обсерватории на телескопах 2,6 м с дисперсией 4 Å/мм и 125 см с дисперсией 8 Å/мм. При расчетах использовались значения $i=40°$, $v_{экв}=210$ км/с, $\varkappa_l=0{,}1989\cdot 10^{-3}$, $\varkappa_d=0{,}468\cdot 10^{-1}$, $c_1=0{,}6$, $c_2=2{,}75\cdot 10^{-3}$; значение коэффициента потемнения к краю для непрерывного спектра u_1 принималось равным 0,6 (подробнее см. [66, 72]). Для звезды с $T_{эфф}=13000°$ К имеем $u_2=0{,}5$. Значение функции $\tau(\Phi, l)$ определялось для 160 областей поверхности. На рис. 40, 41 показаны полученная карта распределения локальных эквивалентных ширин и соответствие наблюдаемых и вычисленных профилей. Погрешность экспериментальных профилей в каждой точке составляет примерно 0,02. Как видно на рис. 40, значение эквивалентной ширины линии меняется на поверхности от 0 до 0,5 Å, что соответствует избытку кремния до 90 раз по сравнению с его нормальным содержанием в звездных атмосферах. Кремний на CUVir локализован в одной области повышенного содержания. Значение параметра регуляризации $\alpha \sim 10^{-9}$.

Метод применялся также для обработки профилей линии λ 4129 Å Eu II в спектре α^2 CVn (HD 112412). Значения угла и экваториальной скорости были выбраны $i=50°$, $v_{экв}=30$ км/сек. Значения $\varkappa_l$, $\varkappa_d$ выбирались такими же, как и в случае звезды CU Vir; $u_1=0{,}6$, $u_2=0$. Параметры зависимости $h(\tau)-c_1$, c_2—были подобраны из условия достаточно точного воспроизведения наблюдаемых профилей (об основах выбора параметров см. в [66]). Они получились равными $c_1=0{,}8$, $c_2=1{,}25\cdot 10^{-2}$. Значения функции $\tau(\Phi, l)$ определялись для 330 областей на поверхности звезды. В результате применения разработанного алгоритма получена карта распределения европия, изображенная на рис. 42. На рис. 43 приведены соответствующие вычисленные и наблюдаемые профили. Концентрация Eu существенно повышена в одной локализованной области на поверхности на долготе от $-50°$ до $+50°$, где она превышает нормальную в 10^6 раз. Однако при этом на долготе около 180° также наблюдается область повышенного содержания Eu. Его обилие в 100 раз превышает нормальное, но значительно меньше полученного в первом пятне.

Среднеквадратичное отклонение вычисленных и наблюдаемых профилей составляет 0,04, т. е. примерно 10% максимума наблюдаемых профилей; значение параметра регуляризации $\alpha \sim 10^{-7}$.

Другие расчеты, полученные с помощью указанного алгоритма, можно найти в [66].

СПИСОК ОБОЗНАЧЕНИЙ

К главам 1, 2

A — точный оператор, действующий из пространства Z в U;
A_h — приближенный оператор;
C — константа алгоритмов, $C>1$;
D — множество ограничений задачи;
$\mathcal{F}$ — семейство допустимых приближенных функционалов (§ 1.1);
$\mathcal{F}_\Delta$ — семейство допустимых приближенных функционалов, отвечающих данному уровню погрешности Δ (§ 1.1);
$\mathcal{F}^m$ — множество сильно возрастающих функций f порядка m;
f — функция, входящая в сглаживающий функционал;
h — погрешность задания оператора;
$I(z)=f[J(z)]$, $I_\delta(z)=f[J_\delta(z)]$;
$I^*=f(J^*)$, $I^*_\delta=f(J^*_\delta)$;
$J(z)$ — точный функционал экстремальной задачи;
$J_\delta(z)$ — приближенный функционал;
$J^*\equiv\inf\{J(z): z\in D\}$, $J^*_\delta\equiv\inf\{J_\delta(z): z\in D\}$;
$M^\alpha[z]$ — сглаживающий функционал;
q — параметр алгоритмов, $q>1$;
u — правая часть операторного уравнения;
u_σ — приближенная правая часть;
Z — пространство решений;
Z^* — множество решений экстремальной задачи, множество квазирешений операторного уравнения;
$\bar{Z}$ — множество Ω-оптимальных решений, квазирешений, псевдорешений;
$Z_0\equiv\{z\in D: \Omega(z)=\Omega^*\}$;
Z^α_δ, Z^α — множество экстремалей сглаживающего функционала при данном α и фиксированном δ;
$\bar{z}$ — Ω-оптимальное решение, квазирешение, псевдорешение;
z^α_δ, z^α — экстремали сглаживающего функционала для данного α;
z^α_+ (z^α_-) — экстремаль сглаживающего функционала, реализующая правое (левое) предельное значение вспомогательных функций;
α, $\alpha(\delta)$, α_δ — параметр регуляризации;
$\alpha_1\equiv\alpha_\delta/q$, $\alpha_2\equiv\alpha_\delta q$;
$\beta(\alpha)$ — вспомогательная функция невязки;
$\gamma(\alpha)$ — вспомогательная функция $\Omega(z^\alpha)$;
δ — вектор погрешности задачи: $\delta=(h,\sigma)$ — для операторных уравнений;
$\varepsilon(\alpha)$ — вспомогательная функция; $\varepsilon(\alpha)=E(z^\alpha)=E_\delta(z^\alpha)$;
$\bar{\varepsilon}(\alpha)=\bar{E}(z^\alpha)$ — модифицированная вспомогательная функция;
$\varepsilon_0\equiv\varepsilon(+0)$, $\varepsilon_\infty\equiv\varepsilon(+\infty)$;
λ_δ — обобщенная мера несовместности: $\lambda_\delta\equiv\inf\{J_\delta(z)+\Psi(\delta,\Omega(z)): z\in D\}$;

$\bar{\mu}$, μ_0 — мера несовместности точного операторного уравнения;
μ_δ — приближенная мера несовместности;
$\nu_\delta \equiv \inf\{I_\delta(z):\ z \in Z_0\}$;
$\pi(\alpha) = \Pi_\delta(z^\alpha)$ — вспомогательная функция;
$\bar{\pi}(\alpha) = \bar{\Pi}_\delta(z^\alpha)$ — модифицированная вспомогательная функция;
$\rho(\alpha) = P_\delta(z^\alpha)$ — вспомогательная функция обобщенной невязки;
$\bar{\rho}(\alpha) = \bar{P}_\delta(z^\alpha)$ — модифицированная обобщенная невязка;
$\rho_0 \equiv \rho(+0)$, $\rho_\infty \equiv \rho(+\infty)$;
σ — погрешность правой части операторного уравнения;
τ — топология в пространстве Z;
$\varphi(\alpha) = M^\alpha[z^\alpha]$ вспомогательная функция;
$\Psi(\delta, \Omega)$, $\Psi_0(\delta, \Omega)$, $\psi(\delta, \Omega)$ — мера аппроксимации функционала;
$\psi(h, \Omega)$ — мера аппроксимации оператора;
$\Omega(z)$ — стабилизирующий (регуляризующий) функционал;
$\Omega_K \equiv \{z \in D:\ \Omega(z) \leqslant K\}$;
$\Omega^* \equiv \inf\{\Omega(z):\ z \in D\}$, $\bar{\Omega} \equiv \inf\{\Omega(z):\ z \in Z^*\}$;
$\bar{\Omega}_1 \equiv \bar{\Omega}(C+q-1)/(C-1)$, $\bar{\Omega}_2 \equiv C\bar{\Omega}/(C-1)$;
ω_δ — параметр алгоритма о.п.к.

К главе 3. Конечномерные аналоги величин

$\hat{D}$ — конечномерный аналог множества D;
H — совокупный вектор погрешности операторного уравнения; $H \equiv (h, 1/M, 1/N)$;
$\hat{I}(\hat{z}) \equiv \hat{I}_\eta(\hat{z}) \equiv f[\hat{J}_\eta(\hat{z})]$;
$\hat{J}_\eta(\hat{z}) \equiv \hat{J}_{\delta N}(\hat{z})$ — конечномерный приближенный функционал;
M — размерность конечномерного пространства U_M;
$\hat{M}^\alpha[\hat{z}]$ — конечномерный сглаживающий функционал;
N — размерность конечномерного пространства Z_N;
P_N, $\bar{P}_N$ — операторы связи пространств Z и Z_N;
$\bar{Q}_M$ — оператор связи пространств U_M и U;
U_M — конечномерное пространство, аппроксимирующее U;
$\hat{u}$, $\hat{u}_{\sigma M}$ — конечномерные правые части операторного уравнения;
Z_N — конечномерное пространство, аппроксимирующее Z;
$\hat{Z}^\alpha$ — множество экстремалей конечномерного сглаживающего функционала при данном α;
$\hat{z}$ — элемент из Z_N;
$\hat{z}^\alpha$ — экстремаль конечномерного сглаживающего функционала;
$\hat{z}_\eta$ — конечномерное приближенное решение;
z_η — приближенное решение, $z_\eta = \bar{P}_N \hat{z}_\eta$;
$\hat{\beta}(\alpha)$, $\hat{\gamma}(\alpha)$, $\hat{\varepsilon}(\alpha)$, $\hat{\bar{\varepsilon}}(\alpha)$, $\hat{\pi}(\alpha)$, $\hat{\bar{\pi}}(\alpha)$, $\hat{\rho}(\alpha)$, $\hat{\bar{\rho}}(\alpha)$ — конечномерные аналоги вспомогательных функций;
η — вектор совокупной погрешности аппроксимации экстремальной задачи:

$$\eta = (\delta, 1/N);$$

$\hat{\lambda} \equiv \lambda_\eta$ — конечномерная обобщенная мера несовместности;
$\hat{\lambda}_\varkappa$ — $\varkappa$-приближение для $\hat{\lambda}$;
$\hat{\mu} \equiv \hat{\mu}_\eta$ — конечномерная приближенная мера несовместности;
$\hat{\nu} \equiv \nu_\eta$ — конечномерный аналог для ν_δ;
$\xi(1/M, \sigma)$ — мера конечномерной аппроксимации правой части операторного уравнения;
$\Psi(\eta, \Omega)$ — совокупная мера аппроксимации задачи;
$\hat{\Psi}(\eta, \Omega)$ — мера конечномерной аппроксимации;
$\hat{\psi}(H, \Omega)$ — мера конечномерной аппроксимации приближенного оператора,

$\hat{\Omega}(\hat{z})$ — конечномерный аналог стабилизирующего функционала;
$\hat{\omega}_\eta$ — параметр конечномерного алгоритма о.п.к.;
$\hat{\Omega}^* \equiv \hat{\Omega}_N^*$ — конечномерный аналог числа Ω^*;

К главе 5

A — матрица размера $m \times n$;
$\bar{A}$ — точная матрица системы линейных уравнений;
A_h — приближенная матрица системы;
$\tilde{A}_h$ — матрица метода м.п.м.;
$\tilde{A}_h^+$ — минимальная псевдообратная матрица;
$\hat{A}_h$ — специальная матрица метода м.п.м.;
$\hat{A}_{\lambda(h)} \equiv A_{\lambda(h)}$ — матрица принципа м.п.м.;
E — единичная матрица размера $m \times m$;
h — погрешность приближенной матрицы;
I — единичная матрица размера $n \times n$;
$\bar{u}$ — точная правая часть системы;
u_σ — приближенная правая часть системы;
$x(\lambda)$ — положительное решение уравнения $x^4 - x^3 = \lambda\rho^{-4}$;
$\bar{z}$ — нормальное псевдорешение системы;
z_δ — приближение к $\bar{z}$;
z_η — приближенное решение по методу м.п.м.;
$\beta(\lambda)$ — невязка метода м.п.м.;
η — вектор погрешностей, $\eta = (h, \sigma)$;
$\theta(x) = \{x^{-1}$ при $x \neq 0$; 0 при $x = 0\}$;
λ, $\lambda(h)$ — параметр регуляризации метода м.п.м.;
$\nu(A)$ — число обусловленности матрицы A;
σ — погрешность задания правой части системы;
$\mathfrak{A}_0$ — множество априорных ограничений на матрицу A;
$\mathfrak{A}$ — нормированное пространство матриц размера $m \times n$;
$\mathfrak{A}^*$ — нормированное пространство матриц размера $n \times m$;
$\mathfrak{A}_n$ — нормированные пространства матриц размера $n \times n$;
$\mathfrak{A}_h = \{A \in \mathfrak{A}_0\colon \; \|A - A_h\| \leqslant h\}$.

СПИСОК ЛИТЕРАТУРЫ

1. Агеев А. Л. Регуляризация нелинейных операторных уравнений на классе разрывных функций // ЖВМ и МФ.— 1980.— Т. 20, № 4.— С. 819—826.
2. Агеев А. Л., Васин В. В. О сходимости обобщенного метода невязки и его дискретных аппроксимаций // Исследования по математическому анализу.— Свердловск: Изд-во УрГУ, 1979.— С. 3—18.
3. Алиев Б. Обобщенный принцип невязки для L-псевдорешений // ЖВМ и МФ.— 1987.— Т. 27, № 2.— С. 286—290.
4. Алифанов О. М., Артюхин Е. А., Румянцев С. В. Экстремальные методы решения некорректных задач.— М.: Наука, 1988.— 288 с.
5. Аллер Л. Распространенность химических элементов.— М.: ИЛ, 1963.
6. Альбер Я. И., Рязанцева И. П. Принцип невязки в нелинейных задачах с монотонными разрывными отображениями — регуляризующий алгоритм // ДАН СССР.— 1978.— Т. 239, № 5.— С. 1017—1020.
7. Альберт А. Регрессия, псевдоинверсия и рекуррентное оценивание.— М.: Наука, 1977.— 223 с.
8. Антипин В. Я. О едином подходе к методам решения некорректных экстремальных задач // Вестник МГУ. Сер. 1, Математика, механика.— 1973. № 2.— С. 61—67.
9. Арсенин В. Я. О разрывных решениях уравнений первого рода // ЖВМ и МФ.— 1965.— Т. 5, № 5.— С. 922—926.
10. Арсенин В. Я. О методах решения некорректно поставленных задач.— М.: Изд. МИФИ, 1977.
11. Арсенин В. Я., Иванов В. В. О решении некоторых интегральных уравнений I рода типа свертки методом регуляризации // ЖВМ и МФ.— 1968.— Т. 8, № 2.— С. 310—321.
12. Баглай Р. Д. О критерии выбора параметра регуляризации, основанном на вычислении функции чувствительности // ЖВМ и МФ.— 1975.— Т. 15, № 2.— С. 305—320.
13. Баев А. В. О построении нормального решения нелинейных некорректных задач методом регуляризации // ЖВМ и МФ.— 1979.— Т. 19, № 3.— С. 594—600.
14. Бакушинский А. Б. Об одном численном методе решения интегральных уравнений Фредгольма I рода // ЖВМ и МФ.— 1965.— Т. 5, № 4.— С. 744—749.
15. Бакушинский А. Б. Один общий прием построения регуляризующих алгоритмов для линейных некорректных уравнений в гильбертовом пространстве // ЖВМ и МФ.— 1967.— Т. 7, № 3.— С. 672—676.
16. Бакушинский А. Б. К распространению принципа невязки // ЖВМ и МФ.— 1970.— Т. 10, № 1.— С. 210—213.
17. Бакушинский А. Б. К принципу итеративной регуляризации // ЖВМ и МФ.— 1979.— Т. 19, № 4.— С. 1040—1043.
18. Бакушинский А. Б. Замечания о выборе параметра регуляризации по критерию квазиоптимальности и отношения // ЖВМ и МФ.— 1984.— Т. 24, № 8.— С. 1258—1259.

19. Бакушинский А. Б., Гончарский А. В. Итеративные методы решения некорректных задач. — М.: Наука, 1989. — 128 с.
20. Бакушинский А. Б., Гончарский А. В. Некорректные задачи. Численные методы и приложения. — М.: Изд-во МГУ, 1989. — 200 с.
21. Бахвалов Н. С. Численные методы. — М.: Наука, 1975. — 632 с.
22. Браун П. А., Киселев А. А. Введение в теорию молекулярных спектров. — Л.: Изд-во ЛГУ, 1983.
23. Будак Б. М., Беркович Е. М. Об аппроксимации экстремальных задач. I // ЖВМ и МФ. — 1971. — Т. 11, № 3. — С. 580—596.
24. Будак Б. М., Виньоли А., Гапоненко Ю. Л. Об одном способе регуляризации экстремальной задачи для непрерывного выпуклого функционала // ДАН СССР. — 1969. — Т. 184, № 1. — С. 12—15.
25. Буша Я. Об одном методе регуляризации решения систем линейных алгебраических уравнений // ДАН СССР. — 1987. — Т. 295, № 1. — С. 11—14.
26. Вайнберг М. М. Вариационный метод и метод монотонных операторов. — М., 1972.
27. Вайникко Г. М. Методы решения линейных некорректно поставленных задач в гильбертовых пространствах. — Тарту: Изд-во Тарт. гос. ун-та, 1982. — 110 с.
28. Вайникко Г. М., Веретенников А. Ю. Итерационные процедуры в некорректных задачах. — М.: Наука, 1986. — 182 с.
29. Вайникко Г. М. О регуляризации некорректных экстремальных задач // Численные методы и оптимизация. — Таллинн: Валгус, 1988. — С. 56—65.
30. Васильев Ф. П. О регуляризации некорректных экстремальных задач // ДАН СССР. — 1978. — Т. 241, № 5. — С. 1001—1004.
31. Васильев Ф. П. Численные методы решения экстремальных задач. — М.: Наука, 1980. — 518 с.
32. Васильев Ф. П. Методы решения экстремальных задач. Задачи минимизации в функциональных пространствах, регуляризация, аппроксимация. — М.: Наука, 1981. — 400 с.
33. Васильев Ф. П., Ковач М. О регуляризации некорректно поставленных экстремальных задач при неточно заданных исходных данных // Computational Math. Banach Center Publ. PWN-Polish Scientific Publishers. Warsaw. — 1984. — V. 13. — С. 237—263.
34. Васильев Ф. П. Оценки скорости сходимости метода регуляризации А. Н. Тихонова для неустойчивых задач минимизации // ДАН СССР. — 1988. — Т. 299, № 4. — С. 792—796.
35. Васин В. В. О β-сходимости проекционного метода для нелинейных операторных уравнений // ЖВМ и МФ. — 1972. — Т. 12, № 2. — С. 492—497.
36. Васин В. В. Метод невязки и конечномерная аппроксимация приближенных решений операторных уравнений // Методы решения условно-корректных задач. — Свердловск: Изд-во УрГУ, 1975. — С. 39—52.
37. Васин В. В. Дискретная сходимость и конечномерная аппроксимация регуляризующих алгоритмов // ЖВМ и МФ. — 1979. — Т. 19, № 1. — С. 11—21.
38. Васин В. В., Танана В. П. Необходимые и достаточные условия сходимости проекционных методов для линейных неустойчивых задач // ДАН СССР. — 1974. — Т. 215, № 5. — С. 1032—1034.
39. Венцель Э. С., Кобылинский В. Г., Левин А. М. Применение метода регуляризации для численного решения задач изгиба тонких упругих пластинок // ЖВМ и МФ. — 1984. — Т. 24, № 2. — С. 323—328.
40. Винокуров В. А. Приближенный метод невязки в нерефлексивных пространствах // ЖВМ и МФ. — 1972. — Т. 12, № 1. — С. 207—212.
41. Винокуров В. А. Два замечания о выборе параметра регуляризации // ЖВМ и МФ. — 1972. — Т. 12, № 2. — С. 481—483.
42. Винокуров В. А. О погрешности приближенного решения линейных обратных задач // ДАН СССР. — 1979. — Т. 11, № 4. — С. 792—793.
43. Винокуров В. П., Гапоненко Ю. Л. Апостериорные оценки решения некорректных обратных задач // ДАН СССР. — 1982. — Т. 263, № 2. — С. 277—280.

44. Воеводин В. В. Вычислительные основы линейной алгебры.— М.: Наука, 1977.— 304 с.
45. Воеводин В. В. О методе регуляризации // ЖВМ и МФ.— 1969.— Т. 9, № 3.— С. 671—673.
46. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления.— М.: Наука, 1984.— 318 с.
47. Гавурин М. К. Лекции по методам вычислений.— М.: Наука, 1971.— 248 с.
48. Гаевский Х., Грегер К., Захариас К. Нелинейные операторные уравнения и операторные дифференциальные уравнения.— М.: Мир, 1978.
49. Гантмахер Ф. Р. Теория матриц.— М.: Наука, 1967.
50. Гапоненко Ю. Л. Об одном регуляризаторе в пространстве непрерывных функций // ЖВМ и МФ.— 1978.— Т. 18, № 2.— С. 379—384.
51. Гапоненко Ю. Л. Метод последовательной аппроксимации для решения нелинейных экстремальных задач // Изв. вузов. Математика.— 1980. № 5.— С. 12—16.
52. Гапоненко Ю. Л. О точности решения нелинейной некорректной задачи при конечном уровне погрешности // ЖВМ и МФ. 1985. Т. 25, № 5.— С. 772—777.
53. Гапоненко Ю. Л. Некорректные задачи на слабых компактах.— М.: Изд-во МГУ, 1989.— 128 с.
54. Гилязов С. Ф. Методы решения линейных некорректных задач.— М.: Изд-во МГУ, 1987.— 120 с.
55. Гилязов С. Ф., Морозов В. А. Оптимальная регуляризация некорректных нормально разрешимых операторных уравнений // ЖВМ и МФ.— 1984.— Т. 24, № 11.— С. 1737—1742.
56. Гласко В. Б., Гущин Г. В., Старостенко В. И. О применении метода регуляризации А. Н. Тихонова к решению нелинейных систем уравнений // ЖВМ и МФ.— 1976.— Т. 16, № 2.— С. 283—292.
57. Гончарский А. В., Гущина Л. Г., Леонов А. С., Ягола А. Г. О некоторых алгоритмах отыскания приближенного решения некорректных задач на множестве монотонных функций // ЖВМ и МФ.— 1972.— Т. 12, № 2.— С. 283—297.
58. Гончарский А. В., Леонов А. С., Ягола А. Г. О решении двумерных уравнений Фредгольма 1-го рода с ядром, зависящим от разности аргументов // ЖВМ и МФ.— 1971.— Т. 11, № 5.— С. 1296—1301.
59. Гончарский А. В., Леонов А. С., Ягола А. Г. Некоторое обобщение принципа невязки для случая оператора, заданного с ошибкой // ДАН СССР.— 1972.— Т. 203, № 6.— С. 1238—1239.
60. Гончарский А. В., Леонов А. С., Ягола А. Г. Об одном регуляризующем алгоритме для некорректно поставленных задач с приближенно заданным оператором // ЖВМ и МФ.— 1972.— Т. 12, № 6.— С. 1592—1594.
61. Гончарский А. В., Леонов А. С., Ягола А. Г. Обобщенный принцип невязки // ЖВМ и МФ.— 1973.— Т. 13, № 2.— С. 294—302.
62. Гончарский А. В., Леонов А. С., Ягола А. Г. Конечно-разностная аппроксимация линейных некорректных задач // ЖВМ и МФ.— 1974.— Т. 14, № 1.— С. 15—24.
63. Гончарский А. В., Леонов А. С., Ягола А. Г. О принципе невязки при решении нелинейных некорректных задач // ДАН СССР.— 1974.— Т. 214, № 3.— С. 499—500.
64. Гончарский А. В., Леонов А. С., Ягола А. Г. О решении некорректно поставленных задач с приближенно заданным оператором // Труды Всесоюзной школы молодых ученых «Методы решения некорректных задач и их применение».— М.: Изд-во МГУ, 1974.— С. 39—43.
65. Гончарский А. В., Леонов А. С., Ягола А. Г. О применимости принципа невязки в случае нелинейных некорректных задач и о новом регуляризующем алгоритме их решения // ЖВМ и МФ.— 1975.— Т. 15, № 2.— С. 290—297.

66. Гончарский А. В., Рябчикова Т. А. и др. Картирование химических элементов на поверхности Ар-звезд. II // Астрон. журн.— 1983.— Т. 60. Вып. 1.— С. 83—90.
67. Гончарский В. А., Степанов В. В. О равномерном приближении решения с ограниченной вариацией некорректных задач // ДАН СССР.— 1979.— Т. 248, № 1.— С. 20—22.
68. Гончарский А. В., Степанов В. В. Численные методы решения некорректных задач на компактных множествах // Вестн. МГУ. Сер. 15.— 1980.— № 3.— С. 12—18.
69. Гончарский А. В., Степанов В. В., Хохлова В. Л., Ягола А. Г. Восстановление локальных профилей линий по наблюдаемым в спектре Ар-звезд // Письма в Астрон. журн.— 1977.— Т. 3, № 6.— С. 278—282.
70. Гончарский А. В., Степанов В. В., Хохлова В. Л., Ягола А. Г. Картирование химических элементов на поверхности Ар-звезд. I // Астрон. журн.— 1982.— Т. 59. Вып. 6.— С. 1146—1156.
71. Гончарский А. В., Черепащук А. М., Ягола А. Г. Численные методы решения обратных задач астрофизики.— М.: Наука, 1978.— 336 с.
72. Гончарский А. В., Черепащук А. М., Ягола А. Г. Некорректные задачи астрофизики.— М.: Наука, 1985.— 352 с.
73. Гончарский А. В., Ягола А. Г. О равномерном приближении монотонного решения некорректных задач // ДАН СССР.— 1969.— Т. 184, № 4.— С. 771—773.
74. Данфорд Н., Шварц Дж. Т. Линейные операторы. Общая теория.— М.: ИЛ, 1962.— 896 с.
75. Дмитриев М. Г., Полещук В. С. О регуляризации одного класса неустойчивых экстремальных задач // ЖВМ и МФ.— 1972.— Т. 12, № 5.— С. 1316—1318.
76. Долгополова Т. Ф., Иванов В. К. О численном дифференцировании // ЖВМ и МФ.— 1966.— Т. 6, № 3.— С. 570—576.
77. Домбровская И. Н. Об уравнениях первого рода с замкнутым оператором // Изв. вузов. Математика.— 1967.— № 6.— С. 39—42.
78. Домбровская И. Н., Иванов В. К. К теории некоторых линейных уравнений в абстрактных банаховых пространствах // Сиб. мат. журн. 1965.— Т. 6, № 3.— С. 499—508.
79. Дэй М. Нормированные линейные пространства.— М.: ИЛ, 1961.
80. Еремин Ю. А., Леонов А. С. О построении устойчивых разностных схем для линейных знаконеопределенных дифференциальных операторов // ЖВМ и МФ.— 1975.— Т. 15, № 3.— С. 635—643.
81. Загонов В. П. Некоторые вариационные методы построения приближений негладких решений некорректно поставленных задач // ЖВМ и МФ.— 1987.— Т. 27, № 11.— С. 1614—1627.
82. Заикин П. Н., Меченов А. С. Некоторые вопросы численного решения интегральных уравнений первого рода методом регуляризации // Научный отчет № 144—ТЗ (468).— М.: ВЦ МГУ, 1971.
83. Зуховицкий С. И., Авдеева Л. И. Линейное и выпуклое программирование.— М.: Наука, 1964.— 348 с.
84. Иванов В. К. Интегральные уравнения первого рода и приближенное решение обратной задачи потенциала // ДАН СССР.— 1962.— Т. 142, № 5.— С. 997—1000.
85. Иванов В. К. О линейных некорректных задачах // ДАН СССР.— 1962.— Т. 145, № 2.— С. 270—272.
86. Иванов В. К. О приближенном решении операторных уравнений первого рода // ЖВМ и МФ.— 1966.— Т. 6, № 6.— С. 1089—1094.
87. Иванов В. К. Некорректные задачи в топологических пространствах // Сиб. мат. журн.— 1969.— Т. 10, № 5.— С. 1065—1074.
88. Иванов В. К., Васин В. В., Танана В. П. Теория линейных некорректных задач и ее приложения.— М.: Наука, 1978.— 206 с.
89. Иоффе В. В., Костиков Р. Р., Разин В. В. Физические методы определения строения органических соединений.— М.: Высшая школа, 1984.
90. Канторович Л. В., Акилов Г. П. Функциональный анализ.— М.: Наука, 1977.— 744 с.

91. Карманов В. Г. Математическое программирование.— М.: Наука, 1975.— 272 с.
92. Каспаров К. Н., Леонов А. С. О фотоэмиссионном методе измерения спектров и новом способе их математической обработки // Журн. прикл. спектроскопии. — 1975. Т. 22, № 3. — С. 491—498.
93. Колмогоров А. Н., Фомин С. Ф. Элементы теории функций и функционального анализа.- М.: Наука, 1968.
94. Коптев Г. С., Пентин Ю. А. Расчет колебаний молекул. М.: Изд-во МГУ, 1977.
95. Коротаев А. Л., Кузнецов П. И., Леонов А. С., Родоман В. Е. Об определении оптимального режима введения антибактериального препарата // Автоматика и телемеханика.— 1986.— № 1.— С. 100—106.
96. Кочетов И. И. О новом способе выбора параметра регуляризации // ЖВМ и МФ.— 1976.— Т. 16, № 2.
97. Кочиков И. В., Курамшина Г. М. Комплекс программ для расчета силовых полей многоатомных молекул по методу регуляризации А. Н. Тихонова // Вест. МГУ. Сер. 2, Химия. — 1985. Т. 26, № 4.— С. 354—358.
98. Кочиков И. В., Курамшина Г. М., Пентин Ю. А., Ягола А. Г. Регуляризующий алгоритм решения обратной колебательной задачи // ДАН СССР.— 1981.— Т. 261, № 5.— С. 1104—1106.
99. Кочиков И. В., Курамшина Г. М., Пентин Ю. А., Ягола А. Г. Об устойчивом решении обратной колебательной задачи // Теория и методы решения некорректных задач.— Новосибирск: ВЦ СО АН СССР, 1983. - С. 124—125.
100. Кочиков И. В., Курамшина Г. М., Пентин Ю. А., Ягола А. Г. Устойчивый метод расчета силовых полей многоатомных молекул в зависимых координатах // Теорет. и эксперим. химия.— 1984.— Т. 20, № 1.— С. 69—75.
101. Кочиков И. В., Курамшина Г. М., Черник С. И., Ягола А. Г. Алгоритм отыскания нормального решения обратной колебательной задачи, основанный на методе Монте-Карло // Вест. МГУ. Сер. 2, Химия.— 1986.— Т. 27, № 6.- С. 597—601.
102. Кочиков И. В., Курамшина Г. М., Ягола А. Г. Устойчивые численные методы решения некоторых обратных задач колебательной спектроскопии // ЖВМ и МФ.— 1987.— Т. 27, № 11.— С. 1651—1661.
103. Кочиков И. В., Матвиенко А. Н., Ягола А. Г. О решении несовместных операторных уравнений с помощью обобщенного принципа невязки // Теория и методы решения некорректно поставленных задач.— Новосибирск: Изд-во НГУ, 1983.— С. 43—48.
104. Кочиков И. В., Матвиенко А. Н., Ягола А. Г. Об одной модификации обобщенного принципа невязки // ЖВМ и МФ.— 1983.— Т. 23, № 6.— С. 1298—1302.
105. Кочиков И. В., Матвиенко А. Н., Ягола А. Г. Обобщенный принцип невязки для решения несовместных уравнений // ЖВМ и МФ.— 1984.— Т. 24, № 7.— С. 1087—1090.
106. Кочиков И. В., Матвиенко А. Н., Ягола А. Г. Метод регуляризации для решения несовместных нелинейных уравнений // ЖВМ и МФ.— 1987.— Т. 27, № 3.— С. 456—458.
107. Крекотень С. П., Леонов А. С., Сухаревский В. Г. Об оптимальном математическом проектировании систем электромагнитов // ДАН СССР.— 1986.— Т. 287, № 2.— С. 312—316.
108. Кряне в А. В., Цупко-Ситников М. В. Нелинейные статистические методы решения плохо обусловленных систем линейных уравнений и их применение к обработке экспериментальных данных // Условно-корректные задачи математической физики и анализа.— Красноярск: Изд-во Красноярского ун-та, 1988.— С. 185—190.
109. Лаврентьев М. М. Об интегральных уравнениях первого рода // ДАН СССР.— 1959.— Т. 127, № 1.— С. 31—33.
110. Лаврентьев М. М. О некоторых некорректных задачах математической физики.— Новосибирск: Наука, 1961.— 92 с.

111. Лаврентьев М. М. Условно-корректные задачи для дифференциальных уравнений.— Новосибирск: Изд-во НГУ, 1973. — 71 с.
112. Лаврентьев М. М., Романов В. Г., Шишатский С. П. Некорректные задачи математической физики и анализа.— М.: Наука, 1980. — 286 с.
113. Ланкастер П. Теория матриц.— М.: Наука, 1982.— 280 с.
114. Ланцош К. Практические методы прикладного анализа.— М.: Физматгиз, 1961.
115. Левин А. М. О регуляризации вычисления нижних граней функционалов // ЖВМ и МФ.— 1984.— Т. 24, № 3.— С. 1123—1128.
116. Леонов А. С. О построении устойчивых разностных схем решения нелинейных краевых задач // ДАН СССР.— 1975.— Т. 224, № 3.— С. 525—528.
117. Леонов А. С. Некоторые аспекты реализации регуляризующего алгоритма обобщенной невязки // Обработка и интерпретация физического элемента.— М.: Изд-во МГУ, 1976.— Вып. 4.— С. 69—81.
118. Леонов А. С. К обоснованию выбора параметра регуляризации по критериям квазиоптимальности и отношения // ЖВМ и МФ.— 1978.— Т. 18, № 6.— С. 1363—1376.
119. Леонов А. С. Об алгоритмах приближенного решения нелинейных некорректных задач с возмущенным оператором // ДАН СССР.— 1979.— Т. 245, № 2.— С. 300—304.
120. Леонов А. С. О выборе параметра регуляризации для нелинейных некорректных задач с приближенно заданным оператором // ЖВМ и МФ.— 1979.— Т. 19, № 6.— С. 1363—1376.
121. Леонов А. С. О регуляризации некорректных задач с разрывными решениями и применение этой методики для решения некоторых нелинейных уравнений // ДАН СССР.— 1980.— Т. 250, № 1.— С. 31—35.
122. Леонов А. С. Кусочно равномерная регуляризация некорректных задач с разрывными решениями // ЖВМ и МФ.— 1982.— Т. 22, № 3.— С. 516—531.
123. Леонов А. С. О связи метода обобщенной невязки и обобщенного принципа невязки для нелинейных некорректных задач // ЖВМ и МФ.— 1982.— Т. 22, № 4.— С. 784—790.
124. Леонов А. С. О применении обобщенного принципа невязки для решения некорректных экстремальных задач // ДАН СССР.— 1982.— Т. 262, № 6.— С. 1306—1310.
125. Леонов А. С. О критериях выбора параметра регуляризации при решении некорректных задач // Методы решения некорректных задач и их приложения.— Новосибирск: Изд-во ВЦ СО АН СССР, 1982.— С. 77—84.
126. Леонов А. С. Метод минимальной псевдообратной матрицы и решение на его основе некорректных задач линейной алгебры // Теория и методы решения некорректно поставленных задач и их приложение.— Новосибирск: Изд-во ВЦ СО АН СССР, 1983.— С. 49—52.
127. Леонов А. С. Применение регуляризующих алгоритмов для решения некоторых задач медицинской радиологии // Теория и методы решения некорректно поставленных задач и их приложения.— Новосибирск: Изд-во ВЦ СО АН СССР, 1983.— С. 132.
128. Леонов А. С. Метод минимальной псевдообратной матрицы для решения некорректных задач линейной алгебры // ДАН СССР.— 1985.— Т. 285, № 1.— С. 36—40.
129. Леонов А. С. Приближенное вычисление псевдообратной матрицы с помощью обобщенного принципа невязки // ЖВМ и МФ.— 1985.— Т. 25, № 6.— С. 133—135.
130. Леонов А. С. О некоторых алгоритмах решения некорректных экстремальных задач // Мат. сб.— 1986.— Т. 129 (171), № 2.— С. 218—231.
131. Леонов А. С. Метод минимальной псевдообратной матрицы // ЖВМ и МФ.— 1987.— Т. 27, № 8.— С. 1123—1138.
132. Леонов А. С. О численной реализации алгоритмов кусочно равномерной регуляризации // ЖВМ и МФ.— 1987.— Т. 27, № 9.— С. 1412—1416.

133. Леонов А. С. Оптимальность по порядку точности обобщенного принципа невязки и некоторых других алгоритмов решения нелинейных некорректных задач с приближенными данными // Сиб. мат. журн. 1988. — Т. XXIX, № 6. С. 85 — 94.
134. Леонов А. С. Численные алгоритмы с неаприорным выбором параметра регуляризации для решения нелинейных некорректных задач // VI Всесоюзный семинар «Обратные задачи и идентификация процессов теплообмена». М.: Изд-во МАИ, 1988. — С. 6 — 7.
135. Леонов А. С. Численные алгоритмы решения нелинейных некорректных задач с неаприорным выбором параметра регуляризации. — М., 1988. — 41 с. — Деп. в ВИНИТИ 15.01.88, № 346 — В — 88.
136. Леонов А. С. Оптимальность по порядку точности регуляризующих алгоритмов для решения нелинейных операторных уравнений // Условно-корректные задачи математической физики и анализа. — Красноярск: Изд-во Красноярского ун-та, 1988. — С. 118 — 124.
137. Леонов А. С., Русин С. П. Оптимальное проектирование концентраторов излучения для целей космической технологии // Труды XXIII чтений, посвященных разработке научного наследия и развитию идей К. Э. Циолковского. Секция «Проблемы космического производства». — М.: Изд-во ИИЕТ АН СССР, 1988. С. 109 — 113.
138. Леонов А. С., Сулейманова М. С. Способ обработки результатов динамических радионуклидных исследований на основе методов решения некорректных задач // Мед. радиология. — 1985. — № 12. — С. 60 — 63.
139. Лисковец О. А. О связи принципа невязки с методом регуляризации // Изв. АН БССР. Сер. физ.-мат. наук. 1972. — № 3. — С. 30 — 34.
140. Лисковец О. А. Метод ε-квазирешений для уравнений 1-го рода // Дифференц. уравнения. — 1973. Т. 9, № 10. С. 1851 1861.
141. Лисковец О. А. Принцип сглаживающего функционала в методе регуляризации для экстремальных задач // ДАН БССР. 1979. — Т. 23, № 1. — С. 16 19.
142. Лисковец О. А. Вариационные методы решения неустойчивых задач. — Минск: Наука и техника, 1981. 344 с.
143. Лоусон Ч., Хенсон Р. Численное решение задачи метода наименьших квадратов. М.: Наука, 1986. 230 с.
144. Люстерник Л. А., Соболев В. И. Краткий курс функционального анализа. — М.: Высшая школа, 1982. — 272 с.
145. Марчук Г. И. Методы вычислительной математики. М.: Наука, 1989. 608 с.
146. Матвеев В. С. Приближенное представление коэффициента поглощения и эквивалентных ширин линий с фойгтовским контуром // Журн. прикл. спектроскопии. — 1972. — Т. 16, № 2. — С. 228 — 233.
147. Мелешко В. И., Задачин В. М. Факторизация и псевдообращение вырожденных возмущенных знаконеопределенных матриц // Изв. вузов. Математика. — 1987, № 11. — С. 42 — 50.
148. Мильман В. Д. Геометрическая теория пространств Банаха. Ч. 2. Геометрия единичной сферы // УМН. — 1971. — Т. 26, № 6. С. 73 149.
149. Морозов В. А. О решении функциональных уравнений методом регуляризации // ДАН СССР. — 1966. — Т. 167, № 3. — С. 510 — 512.
150. Морозов В. А. О регуляризации некорректно поставленных задач и выборе параметра регуляризации // ЖВМ и МФ. 1966. — Т. 6, № 1. — С. 170 — 175.
151. Морозов В. А. Метод регуляризации и некоторые его приложения: Дис. ... канд. физ.-мат. наук. — М., 1967.
152. Морозов В. А. О принципе невязки при решении операторных уравнений методом регуляризации // ЖВМ и МФ. — 1968. — Т. 8, № 2. — С. 295 — 309.
153. Морозов В. А. О регуляризации некоторых классов экстремальных задач // Вычислительные методы и программирование. М.: Изд-во МГУ, 1969. Вып. 12. С. 24 — 37.
154. Морозов В. А. О псевдорешениях // ЖВМ и МФ. 1969. — Т. 9, № 6. — С. 1387 — 1391.

155. Морозов В. А. Об одном новом подходе к решению линейных уравнений 1-го рода с приближенным оператором // Труды 1-й конф. молодых ученых ф-та ВМиК МГУ.— М.: Изд-во МГУ, 1973.— С. 22—28.
156. Морозов В. А. О некоторых общих условиях регуляризуемости некорректных вариационных задач // Труды 1-й конф. молодых ученых ф-та ВМиК МГУ.— М.: Изд-во МГУ, 1973.— С. 140—164.
157. Морозов В. А. Линейные и нелинейные некорректные задачи // Математический анализ (Итоги науки и техники).— М.: ВИНИТИ, 1973.— Т. 11.— С. 129—178.
158. Морозов В. А. О вычислении нижних граней функционалов по приближенной информации // ЖВМ и МФ.— 1973.— Т. 13, № 4.— С. 1045—1048.
159. Морозов В. А. Об принципе невязки при решении несовместных уравнений методом регуляризации А. Н. Тихонова // ЖВМ и МФ.— 1973.— Т. 13, № 5.— С. 1099—1111.
160. Морозов В. А. Регулярные методы решения некорректных задач.— М.: Изд-во МГУ, 1974.—360 с.
161. Морозов В. А. О принципе оптимальности невязки при приближенном решении уравнений с нелинейными операторами // ЖВМ и МФ.— 1974.— Т. 14, № 4.— С. 819—827.
162. Морозов В. А. Регулярные методы решения нелинейных операторных уравнений // Изв. вузов. Математика.— 1978.— № 11.— С. 74—86.
163. Морозов В. А., Гилязов С. Ф. О численном решении некоторых типов несовместных уравнений // Численный анализ на Фортране. Вычислительные методы и инструментальные системы.— М.: Изд-во МГУ, 1979.— С. 69—79.
164. Морозов В. А. Методы регуляризации неустойчивых задач.— М.: Изд-во МГУ, 1987.—216 с.
165. Мудров В. И., Кушко В. Л. Методы обработки измерений.— М.: Радио и связь, 1983.—304 с.
166. Музылев Н. В. Об алгоритме упрощенной регуляризации // ЖВМ и МФ.— 1975.— Т. 15. № 3.— С. 772—775.
167. Мэйндональд Дж. Вычислительные алгоритмы в прикладной статистике.— М.: Финансы и статистика, 1988.—350 с.
168. Натансон И. П. Теория функций вещественной переменной.— М.: Наука, 1974.—480 с.
169. Павлова В. М., Хохлова В. Л. Связь параметров аналитического представления профиля линии в спектре звезды // Науч. информ. Астрон. совета АН СССР.— 1980.— Т. 43.— С. 49—54.
170. Полак Э. Численные методы оптимизации. Единый подход.— М.: Мир, 1974.—376 с.
171. Поляк Б. Т. Методы оптимизации при наличии ограничений // Математический анализ (Итоги науки и техники).— М.: ВИНИТИ, 1974.— Т. 12.— С. 147—198.
172. Пшеничный Б. Н. Метод линеаризации.— М.: Наука, 1983.
173. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах.— М.: Наука, 1975.
174. Русин С. П., Леонов А. С. Об оптимальном математическом проектировании высокотемпературных излучателей // Изв. АН СССР. Энергетика и транспорт.— 1987.— № 4.— С. 154—158.
175. Русин С. П., Леонов А. С. Решение обратных задач теплообмена излучением с учетом зависимости оптических свойств поверхностей от температуры и длины волны // Обратные задачи и идентификация процессов теплообмена.— М.: Изд-во МАИ, 1988.— С. 109—110.
176. Русин С. П., Пелецкий В. Э. Тепловое излучение полостей.— М.: Энергоатомиздат, 1987.—152 с.
177. Самарский А. А. Теория разностных схем.— М.: Наука, 1977.—552 с.
178. Соболев С. Л. Некоторые применения функционального анализа в математической физике.— Л.: Изд-во ЛГУ, 1950.
179. Страхов В. Н. О решении линейных некорректных задач в гильбертовом пространстве // Дифференц. уравнения.— 1970.— Т. 6, № 8.— С. 1490—1495.

180. Страхов В. Н. О построении оптимальных по порядку приближенных решений линейных условно-корректных задач // Дифференц. уравнения.— 1973.— Т. 9, № 10.— С. 1862—1874.
181. Страхов В. Н., Валяшко Г. М. Адаптивная регуляризация линейных некорректных задач // Всесоюзная конференция по некорректно поставленным задачам.— Фрунзе: Изд-во Илим, 1979.— С. 109—110.
182. Страхов В. Н. О сглаживании и трансформации наблюденных значений потенциальных полей методом «дробной» регуляризации // Изв. АН СССР. Физика Земли.— 1988.— № 1.— С. 66—81.
183. Танана В. П. Об одном проекционно-итеративном алгоритме для операторных уравнений первого рода с возмущенным оператором // ДАН СССР.— 1975.— Т. 224, № 5.— С. 1028—1029.
184. Танана В. П. Проекционные методы и конечно-разностная аппроксимация линейных некорректных задач // Сиб. мат. журн.— 1975.— Т. 16, № 2.— С. 1301—1307.
185. Танана В. П. Оптимальные по порядку методы решения нелинейных некорректно поставленных задач // ЖВМ и МФ.— 1976.— Т. 16, № 2.— С. 503—507.
186. Танана В. П. О выборе параметра регуляризации при решении некорректных задач // Всесоюзная конф. по некорректно поставленным задачам. Тезисы докладов.— Фрунзе: Изд-во Илим, 1979.— С. 113—114.
187. Танана В. П. Методы решения операторных уравнений.— М.: Наука, 1981.— 158 с.
188. Танана В. П., Рекант М. А., Янченко С. И. Оптимизация методов решения операторных уравнений.— Свердловск: Изд-во Урал. ун-та, 1987.— 200 с.
189. Тихонов А. Н. Об устойчивости обратных задач // ДАН СССР.— 1943.— Т. 39, № 5.— С. 195—198.
190. Тихонов А. Н. О решении некорректно поставленных задач и методе регуляризации // ДАН СССР.— 1963.— Т. 151, № 3.— С. 501—504.
191. Тихонов А. Н. О регуляризации некорректно поставленных задач // ДАН СССР.— 1963.— Т. 153, № 1.— С. 49—52.
192. Тихонов А. Н. О решении нелинейных интегральных уравнений // ДАН СССР.— 1964.— Т. 156, № 6.— С. 1296—1299.
193. Тихонов А. Н. О нелинейных уравнениях первого рода // ДАН СССР.— 1965.— Т. 161, № 5.— С. 1023—1026.
194. Тихонов А. Н. О некорректных задачах линейной алгебры и устойчивых методах их решения // ДАН СССР.— 1965.— Т. 163, № 3.— С. 591—594.
195. Тихонов А. Н. О методах регуляризации задач оптимального управления // ДАН СССР.— 1965.— Т. 162, № 4.— С. 763—765.
196. Тихонов А. Н. О некорректных задачах оптимального планирования и устойчивых методах их решения // ДАН СССР.— 1965.— Т. 164, № 3.— С. 507—510.
197. Тихонов А. Н. Об устойчивости алгоритмов для решения вырожденных систем линейных алгебраических уравнений // ЖВМ и МФ.— 1965.— Т. 5, № 4.— С. 718—722.
198. Тихонов А. Н. О некорректных задачах оптимального планирования // ЖВМ и МФ.— 1966.— Т. 6, № 1.— С. 81—89.
199. Тихонов А. Н. Об устойчивости задачи оптимизации функционалов // ЖВМ и МФ.— 1966.— Т. 6, № 4.— С. 631—634.
200. Тихонов А. Н. О приближенных системах линейных алгебраических уравнений // ЖВМ и МФ.— 1980.— Т. 20, № 6.— С. 1371—1383.
201. Тихонов А. Н. О задачах с неточно заданной исходной информацией // ДАН СССР.— 1985.— Т. 280, № 3.— С. 559—562.
202. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач.— М.: Наука, 1979.— 285 с.
203. Тихонов А. Н., Васильев Ф. П. Методы решения некорректных экстремальных задач // Math. Models. and Numer. Methods. Banach. Center Publ. Warsaw.— 1978.— V3.— С. 297—342.

204. Тихонов А. Н., Гласко В. Б. О приближенном решении интегральных уравнений Фредгольма I рода // ЖВМ и МФ.- 1964. Т. 4, № 3. - С. 564—571.
205. Тихонов А. Н., Гласко В. Б. Применение метода регуляризации в нелинейных задачах // ЖВМ и МФ.—1965.— Т. 5, № 3.- С. 463—473.
206. Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. Регуляризирующие алгоритмы и априорная информация.— М.: Наука, 1983. - 198 с.
207. Тихонов А. Н., Дмитриев В. И., Чечкин А. В., Березина Н. И. Развитие вариационных методов решения задач синтеза антенн // Вычислительные методы и программирование.— М.: Изд-во МГУ, 1973.— Вып. 20. С. 246 257.
208. Тихонов А. Н., Рютин А. А., Агаян Г. М. Об устойчивом методе решения задачи линейного программирования с приближенными данными // ДАН СССР. - 1983.—Т. 272, № 5. С. 1058 1063.
209. Тихонов А. Н., Уфимцев М. В. Статистическая обработка результатов эксперимента. М.: Изд-во МГУ, 1988. 174 с.
210. Тихонравов А. В. О принципиальной достижимой точности решения задач синтеза // ЖВМ и МФ.— 1982.- Т. 22. № 6.— С. 1421 1433.
211. Треногин В. А. Функциональный анализ. М.: Наука, 1980. 496 с.
212. Уилкинсон Дж., Райнш К. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра. —М.: Машиностроение, 1976.
213. Унзольд А. Физика звездных атмосфер. М.: ИЛ, 1949.
214. Федотов А. М. Линейные некорректные задачи со случайными ошибками в данных.— Новосибирск: Наука, 1982. 190 с.
215. Фиакко А., Мак-Кормик Г. Нелинейное программирование. Методы безусловной последовательной оптимизации. М.: Мир, 1972. 240 с.
216. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений.—М.: Мир, 1980. 280 с.
217. Функциональный анализ (Серия СМБ)/Под ред. С. Г. Крейна.— М.: Наука, 1972.—544 с.
218. Хохлова В. Л. О формализации обратной задачи нахождения локальных профилей по наблюдаемым в спектре Ар-звезд // Astron. Nachr. 1976.— Bd. 297, H. 4. S. 203- 206.
219. Химмельблау Д. Прикладное нелинейное программирование. М.: Мир, 1975. -534 с.
220. Шмульян В. Л. Некоторые геометрические свойства сферы в пространствах типа /B/ // ДАН СССР.—1939. Т. 24, № 7.— С. 648—652.
221. Шолохович В. Ф. Об одном способе численного решения неустойчивых экстремальных задач // Изв. вузов. Математика.—1972.— № 6.— С. 92—100.
222. Ягола А. Г. О выборе параметра регуляризации по обобщенному принципу невязки // ДАН СССР. 1979. Т. 245, № 1. С. 37 39.
223. Ягола А. Г. Обобщенный принцип невязки в рефлексивных пространствах // ДАН СССР. 1979. Т. 249, № 1.— С. 71 —73.
224. Ягола А. Г. О решении нелинейных некорректных задач с помощью обобщенного метода невязки // ДАН СССР.—1980.—Т. 252, № 4.— С. 810—813.
225. Ягола А. Г. О выборе параметра регуляризации при решении некорректных задач в рефлексивных пространствах // ЖВМ и МФ. 1980. Т. 20, № 3.— С. 586—596.
226. Ягола А. Г. Некорректно поставленные задачи с приближенно заданным оператором // Computational Matematics Banach Center Publ. PWN-Polish Scientific Publishers. Warsaw. 1984.— V. 13.— P. 265—279.
227. Björck A., Eldén L. Methods in numerical algebra for ill-posed problems // Preprint LITH-MAT-R-33-1979/Linköping Univ. Dept. of Mathematics. Sweden., 1979.
228. Diestel J. Geometry of Banach spaces-selected topics // Lect. Notes. Math.— 1975. V. 485, N° XI.
229. Hadamard J. Le probleme de Cauchy et les equations aux derivées partielles linearies hyperboliques. - Paris: Hermann, 1932.

230. Hofmann B. Regularisation for Applied Inverse and Ill-Posed Problems.– Leipzig: BSB Teubner, 1986.– 196s.
231. Gauss G. F. Theoria motus corporum coelestium in sectionibus conicis solem ambientium.–– Hamburgi, 1809.
232. Golub G. H. Least squares, singular values and matrix approximations // Aplikace matematiky.— 1968. – V. 13, N° 1.— P. 44—51.
233. Greville T. N. E. Some applications of the pseudoinverse of a matrix // SIAM Rev.— 1960.–– V. 2, N° 1. – P. 15—22.
234. Groetsch C. W. The Theory of Tikhonov Regularization for Fredholm equations of the First Kind.– Boston: Pitman Adv. Publ. Progr., 1984.––104 p.
235. Kochikov I. V., Yagola A. G., Kuramshina G. M. et al. Force fields and mean amplitudes of vibration of chromium, molybdenum and tungsten oxotetrafluorides // J. Molec. Structure.– –1984.— V. 106.— P. 355—360.
236. Kochikov I. V., Yagola A. G., Kuramshina G. M. et al. Calculation of force fields of chromium, molybdenum and tungsten hexafluorides and dioxodifluorides by means of the Tikhonov regularization method // Spectrochim. Acta.––1985.— V. 41A. – N° 1/2. – P. 185—189.
237. Legendere A. M. Novelles methodes pour la determination des orbites des cometes.– Paris: Courcier, 1806.
238. Moore E. N. On the reciprocal of the general algebraic matrix // Bull. Amer. Math. Soc.— 1920. – V. 26.— P. 394—395.
239. Phillips D. L. A technique for numerical solution of certain intergal equations of the first kind // J. Assoc. Comput. Mach.– 1962.– V. 9, N° 1.– P. 84—97.
240. Penrose R. A. A generalized inverse for matices // Proc. Camb. Phil. Soc.– 1955. – V. 51, N° 3.–– P. 406—413.
241. Varah J. M. On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems // SIAM J. Numeral Anal. –1973.–– V. 10. – P. 257—267.
242. Wedin P. Å. Perturbation theory for pseudo-inverses/BIT (Sver).—1973.— V. 13, N°2.– P. 217—232.

Научное издание

ТИХОНОВ Андрей Николаевич
ЛЕОНОВ Александр Сергеевич
ЯГОЛА Анатолий Григорьевич

НЕЛИНЕЙНЫЕ НЕКОРРЕКТНЫЕ ЗАДАЧИ

Редакторы: *И. В. Викторенкова, Л. Г. Полякова*
Художественный редактор *Г. М. Коровина*
Технический редактор *И. Ш. Аксельрод*
Корректоры: *М. А. Смирнов, И. Я. Кришталь*

ИБ № 41224

ЛР № 020297 от 27.11.91. Сдано в набор 19.12.91. Подписано к печати 10.05.95. Формат 60×90/16. Бумага книжно-журнальная. Гарнитура Таймс. Печать офсетная. Усл. печ. л. 19,5. Усл. кр.-отт. 19,5. Уч.-изд.л. 21,39. Тираж 1000 экз. Заказ № 2825. С–021.

Издательская фирма
«Физико-математическая литература» РАН
117071 Москва В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени МПО «Первая Образцовая типография» Министерства печати и информации Российской Федерации. 113054 Москва, Валовая, 28

Отпечатано в Московской типографии № 2 РАН
121099 Москва Г-99, Шубинский пер., 6